Estimation of extremes for heavy-tailed and light-tailed distributions in the presence of random censoring

Julien Worms, Rym Worms

Estimation of extremes for heavy-tailed and light-tailed distributions in the presence of random censoring

Julien Worms (1) & Rym Worms[1] (2)

(1) Université Paris-Saclay/Université de Versailles-Saint-Quentin-En-Yvelines
Laboratoire de Mathématiques de Versailles (CNRS UMR 8100),
F-78035 Versailles Cedex, France,
e-mail : julien.worms@uvsq.fr

(2) Université Paris-Est
Laboratoire d'Analyse et de Mathématiques Appliquées
(CNRS UMR 8050),
UGE, UPEC, F-94010, Créteil, France,
e-mail : rym.worms@u-pec.fr

Abstract In this paper, the flexible semi-parametric model introduced in Gardes et al. (2011) is considered for conducting tail inference of censored data. Both the censored and the censoring variables are supposed to belong to this family of distributions, and thus solutions for modeling the tail of censored data which are between Weibull-tail and Pareto-tail behavior are proposed. Estimators of the tail parameters and extreme quantiles are defined without prior knowledge of censoring strength and asymptotic normality results are proved. Various combinations of the tails of censored and censoring distributions are covered, ranging from rather mild censoring to severe censoring in the tail, *i.e.* when the ultimate probability of censoring in the tail is zero. Finite sample behavior is presented via some simulations and an illustration on real data is also provided.

AMS Classification. Primary 62G32 ; Secondary 62N02

Keywords and phrases. Log-Weibull tail. Tail inference. Random censoring.

[1] Corresponding author

1. Introduction

This paper proposes new contributions to the topic of extreme value statistics for data which are randomly censored from the right.

Consider the classical random censoring setup, where one observes a sample from a couple $(Z, \delta) = (\min(X, C), \mathbb{I}_{X \leqslant C})$ with X denoting the variable of interest, and C a censoring variable (independent from X) which may prevent the user from observing the data X. The observed data is a sample $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ where $(X_1, \ldots, X_n)$ and $(C_1, \ldots, C_n)$ are independent samples of i.i.d. copies of X and C and $\delta_i = \mathbb{I}_{X_i \leqslant C_i}$. The topic of extreme value statistics for randomly censored data deals with the estimation of the tail of X (tail parameters, extreme quantiles, rare probabilities of exceeding a large value), while observing such an incomplete data sample.

A variety of topics can fit this formal random censoring framework. For instance, in the so-called survival analysis setting, the statistical units are patients suffering from a disease, and $X_1, \ldots, X_n$ denote the times elapsed between their inclusion in a study and some event of interest (recovery, recurrence of the disease, death, etc...). Due to loss of follow-up or end of study, one only observe durations $(Z_1, \ldots, Z_n)$ and non-censoring indicators $(\delta_1, \ldots, \delta_n)$. An extreme quantile, in this survival analysis context, is a duration x_p that the "lifetime" of a patient is expected to exceed only with a small probability p, typically smaller than $1/n$. Due to scarcity of data in the tail, some sort of semi-parametric modeling is required to estimate such extreme quantile x_p.

This topic has benefited from a number of contributions in the recent years, which were stimulated by applications in a variety of domains, mainly reliability analysis, survival/lifetime analysis and insurance. Beirlant et al. (2007) and Einmahl et al. (2008) presented a general method for adapting estimators of the extreme value index in this censorship framework. Worms and Worms (2014), Beirlant et al. (2019) and Worms and Worms (2015) proposed a more survival analysis-oriented approach, the first two being restricted to the heavy tail case. Worms and Worms (2018) extended this survival analysis approach to competing risks. The Weibull-tail class of distributions is studied in Worms and Worms (2019). Ndao et al. (2014), Ndao et al. (2016) and Stupfler (2016) extended the framework to data with covariate information. Beirlant et al. (2016), Beirlant et al. (2018) and Goegebeur et al. (2019a) considered the bias-reduction problem. The multivariate case is studied in Goegebeur et al. (2019b) and Hashorva et al. (2014). See also Beirlant et al. (2010) , Gomes and Neves (2011), Sayah et al. (2014), Brahimi et al. (2015), Brahimi et al. (2016), Brahimi et al. (2018), Stupfler (2019) and Bladet al. (2020) for other papers on the subject.

A characteristic of most of these papers is that X and C are always supposed to *share the same type of tail*, *i.e.* a heavy tail censored by a heavy tail, a light (Gumbel) tail censored by another light tail, or a finite tail censored by a finite tail. This is for instance very well described by the 3 cases exhibited in formula (7) of the insightful paper Einmahl et al. (2008).

The main and initial objective of this paper is to broaden the type of tails in the Gumbel domain that the user will be able to deal with, for estimating tail parameters and extreme quantiles based on censored data. As a matter of fact, the lighter-than-Pareto-tails situation was slightly overlooked in censored extremes works, and this may be considered unfortunate since several applications of the censored extremes question do not necessarily exhibit a heavy tail behavior (particularly in survival/lifetime analysis). Essentially only two research papers proposed so far solutions for dealing with light tails. The first one is Einmahl et al. (2008) which proposed estimators of the extreme value index and of extreme quantiles in various cases and in particular in the double Gumbel case (a distribution in the Gumbel domain of attraction censored by another distribution in the same domain). However, the results on the extreme value index, in this case, are stated with a restrictive assumption on the ultimate probability of non-censoring in the tail and there is no formal convergence statement for the proposed extreme quantiles estimator. The second one is Worms and Worms (2019) which considered the general two Weibull-tails framework (a distribution in the Weibull-tail class censored by another distribution in the same class) : this is a strict subset of the double Gumbel case, allowing however interesting configurations where the ultimate probability of non-censoring in the tail can be zero (see its definition in next Section).

The basement of the present work is the flexible semi-parametric model proposed in Gardes et al. (2011) (model $A_1(\tau, \theta)$ described in the next section), which encompasses a large part of the Gumbel maximum domain of attraction and the whole Fréchet one, and therefore provides a more flexible option for modeling various phenomena. In this paper, estimation of the parameters of this model will be made possible in the presence of censoring, with very simple expressions for the estimators. In addition, this setup will allow for a more diverse combination of tails (without prior knowledge of that combination) than the Fréchet versus Fréchet or the Weibull-tail versus Weibull-tail cases (see next section).

The paper is organized as follows. Section 2 formally settles the framework and describes how the parameters of the observed Z can be deduced from those of X and C, thus explaining what is statistically at stake. Section 3 explains how the parameters and extreme quantiles of X can be estimated from the observed censored data, while Section 4 states the main results of this paper, along with the required assumptions on the number k_n of order statistics retained for the estimation. Section 5 contains simulations to illustrate the performance of our estimators and Section 6 an illustration on real-data. Part A to D of the Appendix are devoted to the proofs of our asymptotic results, while part E contains important technical results. Technical aspects of the proofs can be found in a Supplementary Material document provided by the authors (Worms and Worms (2021)).

2. Description of the framework and assumptions

2.1. Model and main assumptions

In the sequel, F^- denotes the general inverse of a function F, $F^-(t) = \inf\{x \in \mathbb{R}; F(x) \geqslant t\}$ for any $t \in [0,1]$.

The formal framework of this paper is the following. Defining for $\tau \in [0,1]$ the Box-Cox function

$$K_\tau(x) = \int_1^x u^{\tau-1} du = \begin{cases} (x^\tau - 1)/\tau & \text{if } \tau \in]0,1], \\ \log(x) & \text{if } \tau = 0, \end{cases}$$

we consider, for parameters $\tau \in [0,1]$ and $\theta > 0$, that a distribution function F belongs to the semi-parametric family $A_1(\tau,\theta)$ if the following holds (see Gardes et al. (2011) where this model was initially introduced in a complete data setting, and El Methni et al. (2012) for additional developments):

$A_1(\tau,\theta)$: for some $x_* > 0$ and every $x \geqslant x_*$, we have

$$1 - F(x) = \exp(-K_\tau^-(\log(H(x)))),$$

where H is an increasing positive function such that H^- is regularly varying at infinity with index θ (which will be denoted by $H^- \in RV_\theta$).

Let us highlight that the tail heaviness of a distribution belonging to $A_1(\tau,\theta)$ is mainly driven by τ, although in practice both shape parameters τ and θ play an important role in the properties and shape of the upper tail. It is easy to see that (for more details see Proposition 2 in Gardes et al. (2011)) :

- $A_1(1,\theta)$ corresponds to distributions in the Fréchet domain of attraction with extreme value index θ (*i.e.* when $1 - F(x) = x^{-1/\theta} l(x)$ with l slowly varying).
- $A_1(0,\theta)$ corresponds to Weibull-tail distributions with Weibull-tail coefficient θ (*i.e.* when $1 - F(x) = \exp(-x^{1/\theta} l(x))$ with l slowly varying).
- The case $\tau \in]0,1[$ corresponds to distributions in the Gumbel domain having tails heavier than Weibull-type ones : such distributions can be conveniently qualified as having log-Weibull-type tails, and log-normal distributions belong to this category with $\tau = 1/2$ (see Gardes et al. (2011) for more examples). The larger τ is, the heavier the tail can be considered.

In this work, the main assumption is that both the censored and the censoring variables have their distribution belonging to the $A_1(\tau,\theta)$ family. This assumption covers a quite flexible setting. Indeed, the users will not need to decide in advance whether a Pareto, Weibull or Log-Weibull tail is convenient for their data, either for the target X or the censoring C. In particular, it is important to note that X and C do not necessarily share the same type of tail. For instance, the cases of Weibull-type data censored by a log-Weibull-type distribution, or of Pareto-type data censored by a log-Weibull-type distribution, are covered. We thus assume the following :

Assumption (A1) : there exist $\tau_X \in [0,1]$, $\tau_C \in [0,1]$, $\theta_X > 0$, $\theta_C > 0$ such that

$$F_X \in A_1(\tau_X,\theta_X) \quad \text{and} \quad F_C \in A_1(\tau_C,\theta_C).$$

This means that there exists positive functions H_X and H_C such that

$$\bar{F}_X(x) = 1 - F_X(x) = \exp(-K_{\tau_X}^{-}(\log(H_X(x)))) \quad \text{and} \quad \bar{F}_C(x) = 1 - F_C(x) = \exp(-K_{\tau_C}^{-}(\log(H_C(x))))$$

and, for some slowly varying functions $\bar{l}_X$ and $\bar{l}_C$ at infinity,

$$H_X^{-}(x) = x^{\theta_X}\bar{l}_X(x) \quad \text{and} \quad H_C^{-}(x) = x^{\theta_C}\bar{l}_C(x).$$

It is clear that under this condition we also have $H_X(x) = x^{1/\theta_X} l_X(x)$ and $H_C(x) = x^{1/\theta_C} l_C(x)$ where both l_X and l_C are slowly varying functions at infinity.

The estimation of the parameters τ_X and θ_X is the main objective of this work (with the estimation of extreme quantiles of F_X being its main application). A first step to do so is to find a relation between the parameters of X and C and those of the observed variable $Z = \min\{X, C\}$.

Under assumption (A1), the following proposition states that the distribution of Z also belongs to the same family of distributions as those of X and C, for some parameters τ_Z and θ_Z specified below :

Proposition 1. *Under Assumption (A1), the distribution function of $Z = \min(X, C)$ satisfies condition $A_1(\tau_Z, \theta_Z)$, where*

$$\tau_Z = \min(\tau_X, \tau_C) \quad \text{and} \quad \theta_Z = \begin{cases} \theta_X & \text{if } 0 \leqslant \tau_X < \tau_C \leqslant 1 \\ \theta_C & \text{if } 0 \leqslant \tau_C < \tau_X \leqslant 1 \\ (\theta_X^{-1/\tau_Z} + \theta_C^{-1/\tau_Z})^{-\tau_Z} & \text{if } 0 < \tau_X = \tau_C \leqslant 1 \\ \min(\theta_X, \theta_C) & \text{if } \tau_X = \tau_C = 0 \end{cases}$$

Therefore, there exists $x_* > 0$ such that for any $x \geqslant x_*$, we have

$$\mathbb{P}(Z > x) = \exp(-K_{\tau_Z}^{-}(\log(H_Z(x)))),$$

where $H_Z^{-} \in RV_{\theta_Z}$. Consequently, if E denotes a standard exponential variable, we have

$$Z \overset{d}{=} H_Z^{-}(\exp K_{\tau_Z}(E)).$$

Remark 1. *It is interesting to note that :*

- *in the two-heavy-tails case $\tau_X = \tau_C = 1$, we recover the well-known fact that $\theta_Z = \gamma_Z = (\gamma_X^{-1} + \gamma_C^{-1})^{-1}$ where γ_X and γ_C are the extreme value indices of X and C (see Beirlant et al. (2007)).*
- *in the two-Weibull-tails case $\tau_X = \tau_C = 0$, we recover the fact that the Weibull-tail parameter of Z is equal to the minimum of those of X and C (see Worms and Worms (2019)).*
- *when $\tau_X = \tau_C$, we have $\theta_Z \leqslant \min(\theta_X, \theta_C)$, but otherwise this is not necessarily the case.*
- *the expression of θ_Z in the fourth case is coherent with the third one in the sense that $\min(\theta_X, \theta_C)$ is indeed the limit of $(\theta_X^{-1/\tau} + \theta_C^{-1/\tau})^{-\tau}$ as $\tau \to 0$.*

In this paper, we will exclude the first two situations evoked in Remark 1 above, which have already been explored in anterior works, and therefore suppose that $(\tau_X, \tau_C) \in [0,1]^2 \backslash \{ (0,0) \, , \, (1,1) \}$.

Let us close this subsection by now describing the more technical assumptions required for our results to hold. This part of the section may be skipped on first reading.

In order to achieve asymptotic normality of the estimators defined in this paper, the slowly varying functions l_X and l_C associated to H_X and H_C are supposed to satisfy a classical second order condition (usually called the SR2 condition) :

Assumption (A2) : there exist some negative constants ρ_X and ρ_C, and some rate functions b_X and b_C having constant sign at $+\infty$ and satisfying $|b_X| \in RV_{\rho_X}$ and $|b_C| \in RV_{\rho_C}$, such that, as $t \to +\infty$,

$$\frac{l_X(tx)/l_X(t) - 1}{b_X(t)} \longrightarrow K_{\rho_X}(x), \quad \text{and} \quad \frac{l_C(tx)/l_C(t) - 1}{b_C(t)} \longrightarrow K_{\rho_C}(x), \forall x > 0. \tag{1}$$

According to the last statement of Proposition 1 and to the expression of our estimators (see next Section), it will be important in the sequel to consider the functions

$$H_Z^{-}(x) = x^{\theta_Z}\tilde{l}(x) \quad \text{and} \quad H_X \circ H_Z^{-}(x) = x^a l(x) \quad \text{with} \quad a := \frac{\theta_Z}{\theta_X}, \tag{2}$$

where both $\tilde{l}$ and l are slowly varying. The crucial parameter $a = \theta_Z/\theta_X$ is equal to 1 in "milde censoring" situations (in particular when $\tau_X < \tau_C$).

In addition, our important technical Lemma 1, stated in Appendix E.1, ensures that functions H_Z^- and $H_X \circ H_Z^-$ also satisfy a second order condition SR2. For technical reasons though, we need to consider the following stronger conditions on $\tilde{l}$ and l, respectively noted $R_{\tilde{l}}(\tilde{b}, \tilde{\rho})$ and $R_l(b, \rho)$, and defined by :

Assumption $\boldsymbol{R_\ell(B,\rho)}$: for some constant $\rho \leqslant 0$ and a rate function B satisfying $\lim_{x\to+\infty} B(x) = 0$, such that for all $\epsilon > 0$, we have

$$\sup_{\lambda \geqslant 1} \left| \frac{\ell(\lambda x)/\ell(x) - 1}{B(x)K_\rho(\lambda)} - 1 \right| \leqslant \epsilon, \text{ for } x \text{ sufficiently large .}$$

Note that, according to Lemma 1 (see Appendix E.1), we have $\rho = \tilde{\rho}$, and that this parameter is negative when either $\tau_X = 0$ or $\tau_C = 0$, but otherwise (*i.e.* in most cases) it is zero, an unpleasant fact which often implies some challenge in the proofs, and affects the rates of convergence of our estimators (with respect to the non-censored framework).

2.2. Proportion of censoring in the tail

It is well known that the strength of censoring affects the statistical performance of estimators in survival analysis. This is naturally also the case for tail estimation under random censoring. Indeed, the ultimate proportion of non-censoring in the tail, denoted p below, explicitly appears in asymptotic variances in the context of extremes of censored data. For instance, the adaptation of the Hill estimator introduced in Einmahl et al. (2008) multiplies the asymptotic variance of the non-censored case by $1/p$ (see Corollary 1 therein), *provided* that this ultimate probability p is positive. Moreover, the pre-asymptotic probability

$$p(x) = \mathbb{P}(\delta = 1 | Z = x)$$

of being non-censored at level x (for large x) often plays a crucial role for proving asymptotic results.

In the context of this paper, Lemma 2 (stated in Appendix E.1) provides precise expansions for $p(x)$, for large x, which turn out to be useful in the proofs of our asymptotic results. In particular, its statement (i) yields the following :

$$\lim_{x\to+\infty} p(x) = p := \begin{cases} 1 & \text{if } 0 \leqslant \tau_X < \tau_C \leqslant 1, \\ 0 & \text{if } 0 \leqslant \tau_C < \tau_X \leqslant 1, \\ \theta_X^{1/\tau_X} / (\theta_X^{1/\tau_X} + \theta_C^{1/\tau_X}) & \text{if } 0 < \tau_X = \tau_C < 1. \end{cases}$$

Note that when X and C are both in the Fréchet or both in the Weibull (*i.e.* finite tail, not Weibull-type tail) maximum domain of attraction, p necessarily belongs to $]0,1[$ (see Einmahl et al. (2008) page 214, for instance). When X and C are both in the Gumbel maximum domain of attraction, things are more complicated, and in this case, Einmahl et al. (2008) considered the assumption $p \in]0,1[$, which is difficult to check in practice and somehow restrictive.

In the model considered in this paper, p can thus span the whole $[0,1]$ interval. In the first situation above (the light censoring one), the fact that the ultimate probability p of non-censoring in the tail is 1 and that the parameters of X are the same as those of Z (see Proposition 1) would suggest that taking into account the censoring is useless. However, as Worms and Worms (2019) already put forward, this is not advisable because those settings produce finite size data where censoring is still present and needs to be taken into account. Similarly, the second situation (strong censoring) where the ultimate probability p is 0 produces, in practice, data which are not completely censored in the tail, and thus the statistical problem of estimating the tail parameters and extreme quantiles of X should and can be addressed. Finally, one can note that the particular situation where tails of X and C have the same heaviness ($\tau_X = \tau_C$) is interesting on its own.

3. Construction of the estimators

Let us denote by Λ_X and Λ_C the cumulative hazard functions associated to F_X and F_C, respectively

$$\Lambda_X(x) = -\log \bar{F}_X(x) \qquad \text{and} \qquad \Lambda_C(x) = -\log \bar{F}_C(x),$$

and let $\hat{\Lambda}_{nX}$ denote the Nelson-Aalen estimator of Λ_X defined as

$$\hat{\Lambda}_{nX}(x) = \sum_{Z_{i,n} \leqslant x} \frac{\delta_{i,n}}{n-i+1}, \tag{3}$$

where $Z_{1,n} \leqslant \ldots \leqslant Z_{n,n}$ are the order statistics of the sample (Z_i) and $\delta_{1,n}, \ldots, \delta_{n,n}$ are the corresponding indicators associated to these reordered Z values. Let $k_n = o(n)$ be an intermediate sequence of integers (which will often be simply denoted by k), representing the number of upper data values retained for tail estimation.

In the following lines, we derive the approximations that inspired our estimators defined below in relation (6). Under assumption (A1), H_X is regularly varying with index $1/\theta_X$ and $K_{\tau_X}(\Lambda_X(x)) = \log(H_X(x))$, hence, for u large, we have

$$K_{\tau_X}(u) \approx \frac{1}{\theta_X} \log(\Lambda_X^{-}(u)).$$

Moreover, for s large and any $u > 1$

$$\log\left(\frac{K_{\tau_X}(su)}{K_{\tau_X}(s)}\right) = \log\left(\frac{(su)^{\tau_X} - 1}{(s)^{\tau_X} - 1}\right) \simeq \tau_X \log u.$$

Combining these two results, we obtain a first approximation, for u and s large, relating τ_X to Λ_X :

$$\tau_X \log u \approx \log\log(\Lambda_X^{-}(su)) - \log\log(\Lambda_X^{-}(s)). \tag{4}$$

The second approximation comes from the fact that, for t large and any given $x > 1$, we have

$$\frac{H_X(tx)}{H_X(t)} = \exp(K_{\tau_X}(\Lambda_X(tx)) - K_{\tau_X}(\Lambda_X(t))) \simeq x^{1/\theta_X},$$

hence θ_X is related to τ_X and Λ_X via the formula :

$$\frac{1}{\theta_X} \log x \approx K_{\tau_X}(\Lambda_X(tx)) - K_{\tau_X}(\Lambda_X(t)). \tag{5}$$

Therefore, applying approximation (4) to $s = \Lambda_X(Z_{n-k_n,n})$ and $u = \Lambda_X(Z_{n-j+1,n})/\Lambda_X(Z_{n-k_n,n})$ on one hand, and approximation (5) to $t = Z_{n-k_n,n}$ and $x = Z_{n-j+1,n}/Z_{n-k_n,n}$ on the other hand, and then plugging in the Nelson-Aalen estimator of Λ_X and summing for $1 \leqslant j \leqslant k$ lead to our proposed estimators of τ_X and θ_X:

$$\hat{\tau}_X := \frac{HH_{k,n}}{D_{k,0}} \quad \text{and} \quad \hat{\theta}_{X,\tau_X} := \frac{H_{k,n}}{D_{k,\tau_X}} \tag{6}$$

with

$$\begin{aligned}
H_{k,n} &:= \frac{1}{k_n} \sum_{j=1}^{k_n} \log(Z_{n-j+1,n}) - \log(Z_{n-k_n,n}), \\
HH_{k,n} &:= \frac{1}{k_n} \sum_{j=1}^{k_n} \log\log(Z_{n-j+1,n}) - \log\log(Z_{n-k_n,n}), \\
D_{k,\tau_X} &:= \frac{1}{k_n} \sum_{j=1}^{k_n} K_{\tau_X}(\hat{\Lambda}_{nX}(Z_{n-j+1,n})) - K_{\tau_X}(\hat{\Lambda}_{nX}(Z_{n-k_n,n})).
\end{aligned}$$

The two estimators above are thus ratios involving on one hand the mean of either the log-spacings (*i.e.* the Hill statistic) or the log-log-spacings, and on the other hand a denominator involving the Nelson-Aalen estimator at the k upper values of the observed Z sequence.

Note that the expressions of the estimators defined in (6) do not depend on the relative positions of τ_X and τ_C (or of θ_X and θ_C). They can be calculated whatever the combinations of the tails of X and C are, with the same formulas. However, we will see in the next Section that the rates of convergence, performances, and assumptions of these estimators can differ depending on the strength of censoring.

Remark 2. *In the case $\tau_X = \tau_C = 0$, corresponding to the purely Weibull-tail framework, the estimator $\hat{\theta}_{X,0}$ corresponds to the one studied in Worms and Worms (2019), because $K_{\tau_Z}(x) = \log(x)$ in that case. In the case $\tau_X = \tau_C = 1$, corresponding to the purely heavy-tail framework, the estimator $\hat{\theta}_{X,1}$ corresponds to the adapted Hill estimator studied in Beirlant et al. (2007), because in that case $K_{\tau_Z}(x) = x - 1$ and thus* *we have exactly $D_{k,1} = \hat{p}_k$ (see formula (7) below). As said earlier, these two particular cases are excluded from the scope of the statements of this paper because properties of $\hat{\theta}_{X,0}$ and $\hat{\theta}_{X,1}$ are already known.*

The main issue in the proofs lies in the treatment of the denominators D_{k,τ_X} and $D_{k,0}$. In fact, the statistic D_{k,τ_X} defined below (6) turns out to be related to the proportion $\hat{p}_k$ of uncensored data in the tail via the relation (see Lemma 3 in the Appendix for the details)

$$D_{k,\tau_X} \approx \left(\hat{\Lambda}_{nX}(Z_{n-k,n})\right)^{\tau_X - 1} \hat{p}_k \quad \text{where} \quad \hat{p}_k := \frac{1}{k}\sum_{j=1}^{k} \delta_{n-j+1,n}$$

because of the nature of the Box-Cox transformation K_{τ_X}, Taylor's formula, and of the fact that

$$\frac{1}{k}\sum_{j=1}^{k}\left(\hat{\Lambda}_{nX}(Z_{n-j+1,n}) - \hat{\Lambda}_{nX}(Z_{n-k,n})\right) = \frac{1}{k}\sum_{j=1}^{k}\sum_{l=j}^{k}\frac{\delta_{n-l+1,n}}{l} = \frac{1}{k}\sum_{j=1}^{k}\delta_{n-j+1,n}. \tag{7}$$

Therefore, the properties of our estimators will rely on a careful study of two sequences. The first one is $\hat{\Lambda}_{nX}(Z_{n-k,n})$ (in particular, how it can be approximated by $\Lambda_X(Z_{n-k,n})$ and written as an increasing function of $\log n/k$; see Lemma 4 in the Appendix). The second one is the sequence $\hat{p}_k$, which converges to 0, 1 or a value $p \in]0,1[$ depending on the position of τ_X with respect to τ_C (Proposition 2 in section Appendix A provides the full details about this, and relies on sharp second order expansions of the different regularly varying functions that appear in this framework, *cf* the important technical Lemmas 1 and 2 in the Appendix).

Finally, let us deal with the estimation of an extreme quantile $x_{p_n} := \bar{F}_X^{-}(p_n)$ of the distribution of X, with $p_n \to 0$, as $n \to +\infty$. Applying the approximation (5) now to $t = Z_{n-k,n}$ and $x = x_{p_n}/Z_{n-k,n}$, we can propose the following estimator of x_{p_n} (with both θ_X and τ_X being unknown) :

$$\hat{x}_{p_n} := Z_{n-k,n} \exp\left\{\hat{\theta}_{X,\hat{\tau}_X}\left(K_{\hat{\tau}_X}(-\log(p_n)) - K_{\hat{\tau}_X}(\hat{\Lambda}_{nX}(Z_{n-k,n}))\right)\right\}. \tag{8}$$

Note that if we know that $\tau_X = 0$ and we then set $\hat{\tau}_X = 0$, then this estimator is the same as the one proposed in Worms and Worms (2019).

4. Asymptotic results

The main assumptions on the model and the different notations have been stated in the previous sections. In order to obtain the asymptotic normality of our estimators, we naturally need the sequence (k_n) (number of top order statistics to use) to satisfy some conditions (we will note $k = k_n$ from now on). The first one is standard in the literature on Weibull-tail or log-Weibull-tail models :

H_1 : $k \to +\infty$, $\frac{k}{n} \to 0$, $\frac{\log k}{\log n} \to 0$, as $n \to +\infty$.

Moreover, introducing the important notation

$$L_{nk} = \log(n/k),$$

let v_n be a factor which will contribute to the rates of convergence of our estimators, and which depends on the censoring strength in the tail :

$$v_n := \begin{cases} 1 & \text{if } 0 < \tau_X < \tau_C \leqslant 1 \ \text{ or } \ 0 < \tau_X = \tau_C < 1 \ \text{ or } \ 0 = \tau_X < \tau_C \leqslant 1, \\ L_{nk}^{\frac{1}{2}(\frac{\tau_C}{\tau_X}-1)} & \text{if } 0 < \tau_C < \tau_X \leqslant 1, \\ L_{nk}^{-1/2}(\log L_{nk})^{\frac{1}{2}(\frac{1}{\tau_X}-1)} & \text{if } 0 = \tau_C < \tau_X \leqslant 1. \end{cases}$$

Note that $v_n = 1$ in the mild or moderate censoring cases (when $p = 1$ or at least $p > 0$, see Section 2.2), and $v_n \to 0$ in the strong censoring cases (when $p = 0$). We also consider the following conditions

$$H_2 : 0 < \tau_X < \tau_C \leqslant 1 \ \text{ and } \begin{cases} (i)\sqrt{k}L_{nk}^{\tau_X/\tau_C - 1} \to 0 \text{ if } \frac{1}{\tau_C} - \frac{1}{\tau_X} \geqslant -1 \\ (ii)\sqrt{k}L_{nk}^{-\tau_X} \to 0 \text{ if } \frac{1}{\tau_C} - \frac{1}{\tau_X} < -1 \end{cases}$$

$$H_3 : 0 < \tau_C < \tau_X \leqslant 1 \ \text{ and } \begin{cases} (i)\sqrt{k}v_n \to +\infty \\ (ii)\sqrt{k}v_n L_{nk}^{\tau_C/\tau_X - 1} \to 0 \text{ if } \frac{1}{\tau_X} - \frac{1}{\tau_C} \geqslant -1 \\ (iii)\sqrt{k}v_n L_{nk}^{-\tau_C} \to 0 \text{ if } \frac{1}{\tau_X} - \frac{1}{\tau_C} < -1 \end{cases}$$

$$H_4 : 0 < \tau_X = \tau_C < 1 \quad \text{and} \quad \sqrt{k} L_{nk}^{-\tau_X} \to 0$$
$$H_5 : 0 = \tau_X < \tau_C \leqslant 1 \quad \text{and} \quad \exists \delta > 0, \sqrt{k} L_{nk}^{\tilde{\rho}+\delta} \to 0$$
$$H_6 : 0 = \tau_C < \tau_X \leqslant 1 \quad \text{and} \quad \begin{cases} (i)\sqrt{k} v_n \to +\infty \\ (ii)\sqrt{k} v_n (\log L_{nk})^{-1} \to 0 \end{cases}$$

(in assumption H_5 above, $\tilde{\rho}$ denotes the second order parameter associated to the slowly varying function $\tilde{l}$, which is negative in this case ; see formula (2) in Section 2 as well as Lemma 1 in Appendix E.1)

Remark 3. *A possible choice of the sequence (k_n) satisfying the conditions above, as well as the additional conditions in the theorems stated below, is :*

$$k_n = \begin{cases} (\log n)^a & \text{if} \quad \tau_C \neq 0, \\ \frac{\log n}{(\log \log n)^b} & \text{if} \quad \tau_C = 0. \end{cases}$$

The choice and scope of exponents a and b depend on which condition $H_2, \ldots,$ or H_6 is considered. Note that these sequences tend to infinity a bit more slowly than the sequences (k_n) considered in El Methni et al. (2012), in the non-censoring situation (see the paragraph following the statement of their Theorem 2).

The following four theorems respectively state the convergence in distribution of the estimators $\hat{\theta}_{X,\tau_X}$ (with τ_X known), $\hat{\tau}_X$, $\hat{\theta}_{X,\hat{\tau}_X}$, and $\hat{x}_{p_n}$, all of them being defined in the previous section.

Theorem 1. *Let assumptions (A1) and (A2) hold, with $(\tau_X, \tau_C) \in [0,1]^2 \backslash \{ (0,0) \, , \, (1,1) \}$, as well as $R_l(b, \rho)$ and $R_{\tilde{l}}(\tilde{b}, \tilde{\rho})$. If (k_n) satisfies H_1 and one of the conditions $H_2, \ldots, H_6$, then we have, as $n \to \infty$,*

$$\sqrt{k} v_n (\hat{\theta}_{X,\tau_X} - \theta_X) \xrightarrow{d} N\left(0, \theta_X^2 \sigma^2\right),$$

where $a = \theta_Z / \theta_X$ and

$$\sigma^2 = \begin{cases} 1 & \text{if } 0 \leqslant \tau_X < \tau_C \leqslant 1, \\ a^{-1/\tau_X} \left(\frac{\tau_X}{\tau_C}\right)^{1-1/\tau_X} & \text{if } 0 < \tau_C < \tau_X \leqslant 1, \\ a^{-1/\tau_X} & \text{if } 0 < \tau_X = \tau_C < 1, \\ a^{-1/\tau_X} \tau_X^{1-1/\tau_X} & \text{if } 0 = \tau_C < \tau_X \leqslant 1. \end{cases}$$

Remark 4. *When $\tau_X < \tau_C$, the ultimate probability p of non-censoring is 1, this is the mild censoring situation. When $\tau_X = \tau_C$, it is easy to see that $\theta_Z < \theta_X$ and thus the asymptotic variance is larger than in the case $\tau_X < \tau_C$ (i.e. we have $\sigma^2 > 1$). When $0 < \tau_C < \tau_X$ (strong censoring setting), the ultimate probability of non-censoring p is zero, and the factor σ^2 is < 1 when $\theta_C > \theta_X$, but otherwise this is not necessarily the case.*

Theorem 2. *Under the same assumptions as Theorem 1, we have, as $n \to \infty$,*

$$\begin{aligned} &\text{if } \tau_X \neq 0, \quad \sqrt{k} v_n (\hat{\tau}_X - \tau_X) \xrightarrow{d} N\left(0, \tau_X^2 \sigma^2\right), \\ &\text{if } \tau_X = 0, \quad \hat{\tau}_X = O_{\mathbb{P}}(1/\log(L_{nk})) \xrightarrow{\mathbb{P}} 0, \end{aligned}$$

Theorem 3. *Under the same assumptions as Theorem 1, if $\tau_X > 0$ and if we further assume that*

$$\frac{\sqrt{k} v_n}{\log L_{nk}} \to +\infty \quad (\text{if } \tau_C \neq 0) \quad \text{or} \quad \frac{\sqrt{k} v_n}{\log \log L_{nk}} \to +\infty \quad (\text{if } \tau_C = 0), \tag{9}$$

we then have, as $n \to \infty$,

$$\begin{aligned} &\text{if } \tau_C \neq 0 \quad \frac{\sqrt{k} v_n}{\log L_{nk}} (\hat{\theta}_{X,\hat{\tau}_X} - \theta_X) \xrightarrow{d} N\left(0, \theta_X^2 \sigma^2 \tau_Z^2\right), \\ &\text{if } \tau_C = 0 \quad \frac{\sqrt{k} v_n}{\log \log L_{nk}} (\hat{\theta}_{X,\hat{\tau}_X} - \theta_X) \xrightarrow{d} N\left(0, \theta_X^2 \sigma^2\right). \end{aligned}$$

Remark 5. *Note that the rate of convergence and asymptotic variance of $\hat{\theta}_{X,\hat{\tau}_X}$ are altered and different from that of $\hat{\theta}_{X,\tau_X}$ due to the plug-in of $\hat{\tau}_X$.*

Theorem 4. *Under the same assumptions as Theorem 3, if moreover*

$$\frac{\sqrt{k}v_n}{\log\log(1/p_n)(-\log(p_n))^{\tau_X}} \to +\infty \tag{10}$$

and

$$\frac{\log L_{nk}}{\log\log(1/p_n)} \to 0 \quad (\textit{if } \tau_C \neq 0) \qquad \textit{or} \qquad \frac{\log\log L_{nk}}{\log\log(1/p_n)} \to 0 \quad (\textit{if } \tau_C = 0), \tag{11}$$

we then have, as $n \to \infty$,

$$\frac{\sqrt{k}v_n}{\log\log(1/p_n)(-\log(p_n))^{\tau_X}}\left(\frac{\hat{x}_{p_n}}{x_{p_n}} - 1\right) \xrightarrow{d} N\left(0, \theta_X^2\sigma^2\right).$$

Remark 6. *Note that condition* (11) *allows for an order p_n of the high quantile which is lower than $\frac{1}{n}$, making it really extreme, while condition* (10) *is a restriction on this order.*

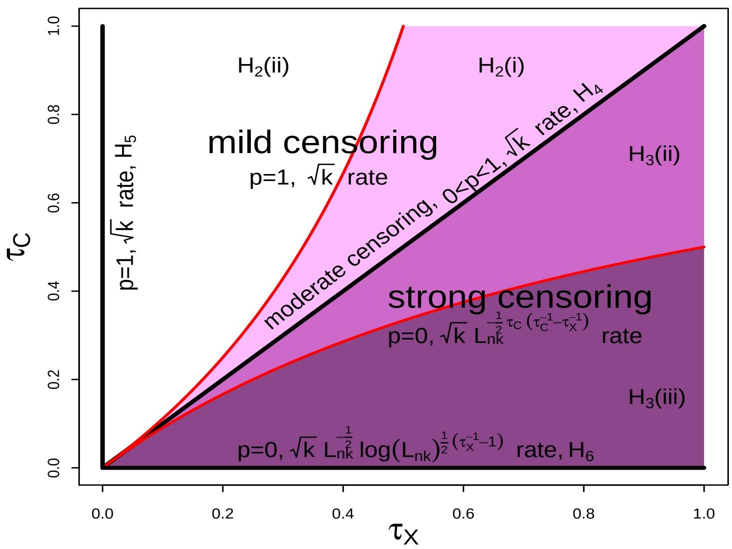


Figure 1: Illustration of the possible combinations of τ-parameters and the impact on rates and assumptions ($\tau_X = \tau_C = 0$ and $\tau_X = \tau_C = 1$ are excluded).

Remark 7. *There is some sort of phase transition phenomenon in the above results. As a matter of fact, not only the rate of convergence of our estimators vary whether τ_X is $\leqslant \tau_C$ or not, but the closeness of the parameters τ_X and τ_C also play a role (see assumptions H_2 and H_3) : the assumptions vary whether τ_X is lower than τ_C but not too close to it (i.e. $1 < \frac{1}{\tau_X} - \frac{1}{\tau_C}$), lower than τ_C but close to it (i.e. $0 < \frac{1}{\tau_X} - \frac{1}{\tau_C} \leqslant 1$), equal to τ_C, larger than and close to τ_C (i.e. $0 < \frac{1}{\tau_C} - \frac{1}{\tau_X} \leqslant 1$), or sufficiently larger than τ_C (i.e. $1 < \frac{1}{\tau_C} - \frac{1}{\tau_X}$). Figure 1 helps to understand these facts.*

However, in practice, for finite and moderate values of n, visualizing these findings on simulations is not easy, because other factors (than just the tail parameters) play a non-negligible role in the estimation quality.

Let us close this section by providing a hint of the proof of the consistency of our estimators (consistency alone is not considered in the full proofs, only asymptotic normality is detailed). Let us note $\hat{\theta}_Z^{(c)}$ and $\hat{\tau}_Z^{(c)}$ the following estimators of θ_Z and τ_Z

$$\hat{\theta}_Z^{(c)} = \frac{H_{k,n}}{\mu_{1,\tau_Z}(L_{nk})} \quad \text{and} \quad \hat{\tau}_Z^{(c)} = \frac{HH_{k,n}}{\mu_{1,0}(L_{nk})} \quad \text{where} \quad \mu_{1,\tau}(t) = \int_0^\infty \left(K_\tau(x+t) - K_\tau(t)\right) e^{-x}\,dx. \tag{12}$$

The first one was introduced in Gardes et al. (2011). The second one is similar to the estimator proposed in Albert et al. (2020) (in a slightly different setting). Using the material of Gardes et al. (2011) and Albert et al. (2020), one can prove that $\hat{\theta}_Z^{(c)}$ and $\hat{\tau}_Z^{(c)}$ are consistent estimators of θ_Z and τ_Z. Our estimators can then be written as

$$\hat{\theta}_{X,\tau_X} = \hat{\theta}_Z^{(c)} \times \frac{\mu_{1,\tau_Z}(L_{nk})}{D_{k,\tau_X}} \quad \text{and} \quad \hat{\tau}_X = \hat{\tau}_Z^{(c)} \times \frac{\mu_{1,0}(L_{nk})}{l\mu_{1,\tau_Z}(L_{nk})} \times \frac{l\mu_{1,\tau_Z}(L_{nk})}{D_{k,0}} \tag{13}$$

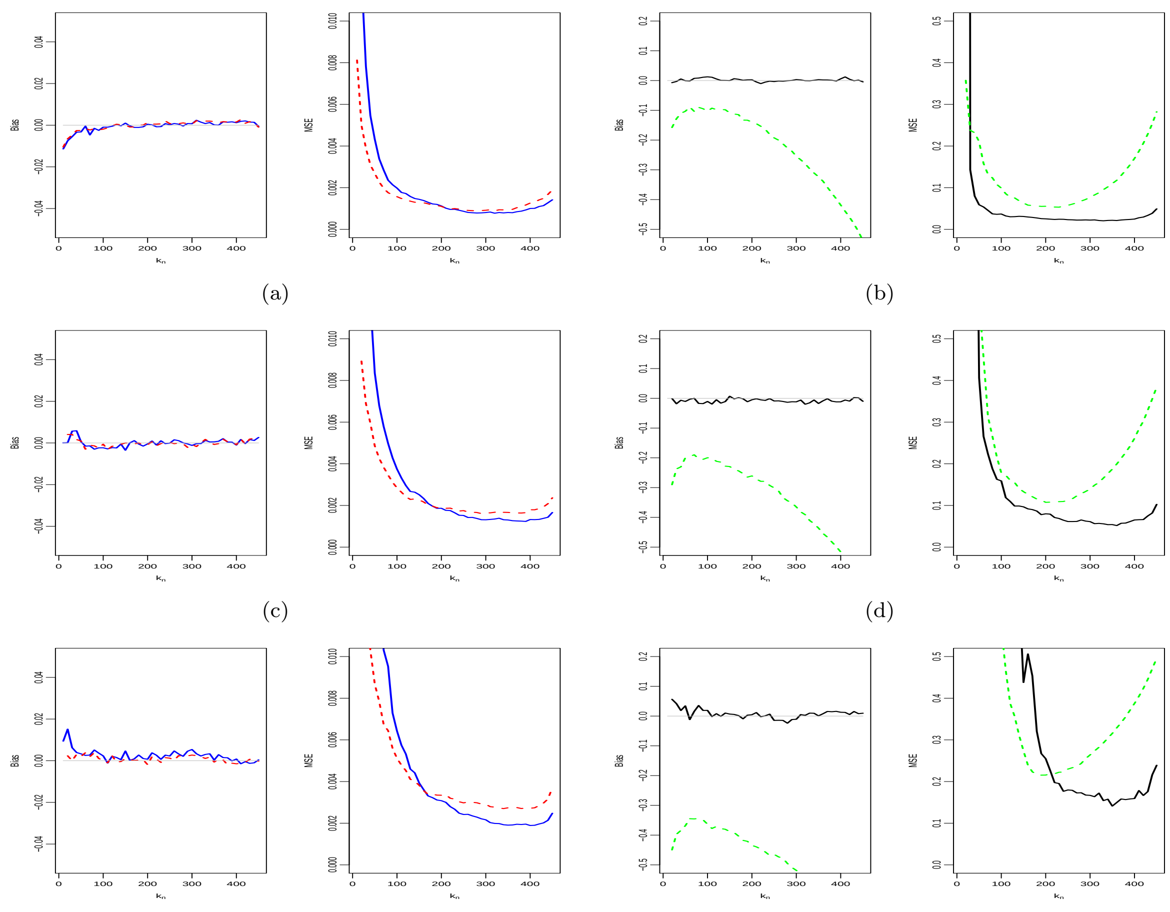
where $l\mu_{1,\tau}(t) := \int_0^\infty (\log(K_\tau(x+t)) - \log(K_\tau(t)))\, e^{-x}\, dx$.

The consistency of $\hat{\theta}_{X,\tau_X}$ will thus come from the convergence of the ratio $\mu_{1,\tau_Z}(L_{nk})/D_{k,\tau_X}$ to $1/a = \theta_X/\theta_Z$, which is deduced from Corollary 1 (stated in Appendix A) of the present paper. The consistency of $\hat{\tau}_X$ comes from the convergence of $l\mu_{1,\tau_Z}(L_{nk})/D_{k,0}$ to τ_X, which is deduced from Corollary 2 (stated in Appendix B), and from the fact that $\mu_{1,0}(t)/l\mu_{1,\tau_Z}(t)$ converges to $1/\tau_Z$ as $t \to \infty$ (which is deduced from relations (A.3) and (B.3) in the Appendix).

It is noteworthy that equation (13) describes a way of adapting to the censoring context any estimators of θ or τ valid in the complete data setting, by simply dividing by the appropriate expression involving D_{k,τ_X} or $D_{k,0}$.

Finally, note that $\hat{\tau}_Z^{(c)}$ defined above is a new estimator of τ in the $A_1(\tau,\theta)$ model without censoring, and thus a competitor of the estimator which was proposed in El Methni et al. (2012) (which required the delicate choice of two intermediate sequences k_n and k_n').

5. Finite sample comparisons



(a) (b) (c) (d) (e) (f)

Figure 2: Simulation with X log-Weibull censored by C log-Normal, where $\tau_X = 0.4 < \tau_C = 0.5$ in first line (figures (a)-(b), mild censoring $p = 1$), $\tau_X = 0.5 = \tau_C$ in second line (figures (c)-(d), moderate censoring $p \in]0,1[$), and $\tau_X = 0.6 > \tau_C = 0.5$ in third line (figures (e)-(f), strong censoring $p = 0$). The graphs represent observed bias and MSE of estimators $\hat{\tau}_X$ (blue) and $\hat{\theta}_{X,\hat{\tau}_X}$ (dashed red) in figures (a)-(c)-(e), and relative bias and MSE of estimators $\hat{x}_{p_n}$ (black) and $\hat{x}_{p_n}^{EFG}$ (dashed green) in figures (b)-(d)-(f).

In this section, we illustrate, using few simulations, the finite sample performances of our estimators of τ_X, θ_X and x_{p_n} (for small p_n), in terms of observed bias and mean squared error (MSE). Note that numerous different situations could be considered with our flexible framework : a thorough and extensive simulation study is however not possible within the limits of the present paper. We generate $N = 1000$ samples of size $n = 500$.

We consider three classes of distributions of Log-Weibull-tail type, for the target variable X and the censoring variable C (for the first two classes, see Proposition 3 in Gardes et al. (2011) for the justification that they fit our framework) :

- Log-Weibull(θ) distribution such that its logarithm has c.d.f. $1 - \exp(-x^{1/\theta})$ $(x > 0)$. It satisfies assumption $A_1(\theta, \theta)$.
- Log-Normal distribution $LN(\mu, \sigma^2)$, which satisfies assumption $A_1(\frac{1}{2}, \frac{\sigma\sqrt{2}}{2})$.
- Model $\mathscr{F}$ with c.d.f. F_τ satisfying $A_1(\tau, 1/5)$, with $H^-(x) = x^{1/5}(1 + x^{-1/2})$ $(\forall x)$.

We then consider three cases : a Log-Weibull(θ_X) distribution censored by the Log-Normal$(1, 1/2)$ distribution (Figure 2), the Log-Normal$(1, 1/2)$ distribution censored by a Log-Weibull(θ_C) distribution (Figure 3), and then a distribution in the $\mathscr{F}$ model censored by another distribution in the $\mathscr{F}$ model (Figure 4). In each case, we consider three situations with $\tau_X < \tau_C$, $\tau_X = \tau_C$ or $\tau_X > \tau_C$, corresponding to different (ultimate) intensities of censoring in the tail.

In parts (a),(c),(e) of Figures 2, 3 and 4, we present the bias and the MSE of our estimators $\hat{\tau}_X$ and $\hat{\theta}_{X,\hat{\tau}_X}$ as a function of k. In parts (b),(d),(f) of Figures 2, 3 and 4, we present the relative bias and the relative MSE of our estimator $\hat{x}_{p_n}$ for the value $p_n = 0.001$, compared with those of the existing estimator defined, in a more general censored setting, by equation (8) in Einmahl et al. (2008) :

$$\hat{x}_{p_n}^{EFG} = Z_{n-k,n} + \hat{a}_k \frac{((1 - \hat{F}_n(Z_{n-k}))/p_n)^{\hat{\gamma}^{c,Mom}} - 1}{\hat{\gamma}^{c,Mom}}, \tag{14}$$

where $\hat{\gamma}^{c,Mom}$ is the moment estimator of the extreme value index γ_X of F adapted to censoring and $\hat{F}_n$ stands for the Kaplan-Meier estimator of the c.d.f. F. We refer to Einmahl et al. (2008) for the expression of $\hat{a}_k$. Note that no formal asymptotic result is currently available for $\hat{x}_{p_n}^{EFG}$.

Concerning the performance of the estimators $\hat{\theta}_{X,\hat{\tau}_X}$ and $\hat{\tau}_X$, we observe that when X has a Log-Weibull tail, the bias and the MSE for both estimators are very small. When one deviates from this situation, though, they are not very satisfactory on the situations presented here. Note however that these estimators are the first to be proposed in this context, which is why no comparison to competitors is presented . Another remark is that the quality of the estimators do not systematically deteriorate when censoring gets stronger.

Concerning the performance of the high quantile estimator, the figures show very good performances when X has a Log-Weibull tail. When one deviates from this situation, things may become worse. It is particularly true here in the Log-Normal versus Log-Weibull case. However, our estimator remains competitive in terms of bias and MSE in a number of situations, for instance in Figure 4.

6. Real data analysis

In this section, we apply our methodology to one of the datasets included in the Cancer Genome Atlas (TCGA, more information at cancergenome.nih.gov), namely the dataset concerning women suffering from an ovarian cancer. The `ov.clinical` dataset, accessible using the R package RTCGA (or manually), contains informations about 574 women[2]. In this section we will only be interested in the vital status of the patient ($\delta = 1$ if the patient died during the study, or $\delta = 0$ if the patient was still alive at the end of the study or was lost in the follow-up), and the observed duration $Z = \min(X, C)$ ($Z = X$ if survival time is actually observed, and $Z = C$ if censoring occurred). The tail of the survival time distribution F_X will be estimated using the model (A1), and in this context an extreme quantile x_p associated to a small probability p, is a survival time that a patient is expected to exceed only with probability p.

We plot, in figure 5, the proportion $\hat{p}_k$ of non-censoring as a function of k_n. We observe that the censoring is rather strong, overall censoring rate is around 40% and more around 55% in the tail. If we consider the

[2]the original dataset contains 591 data lines, but 17 of them were not workable because of missing vital status or missing survival time.

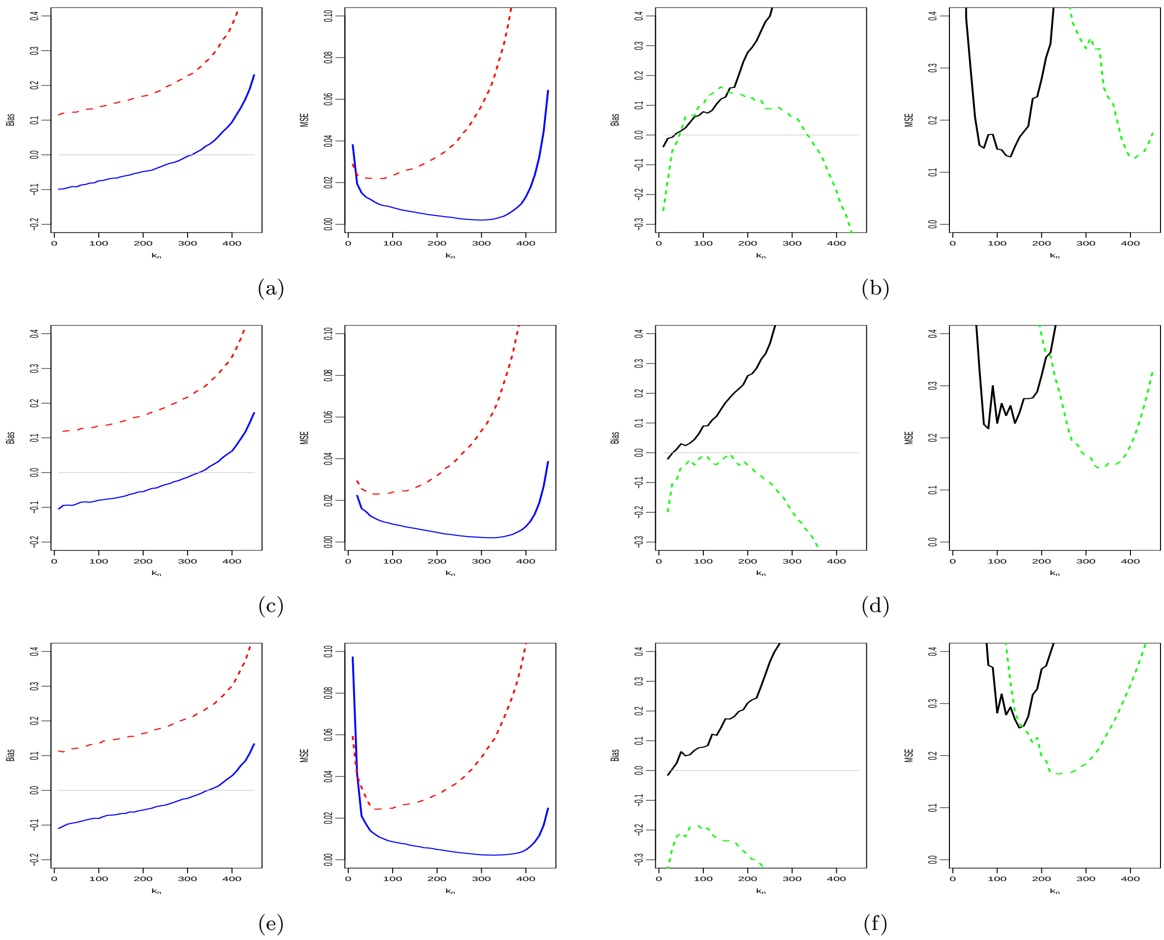
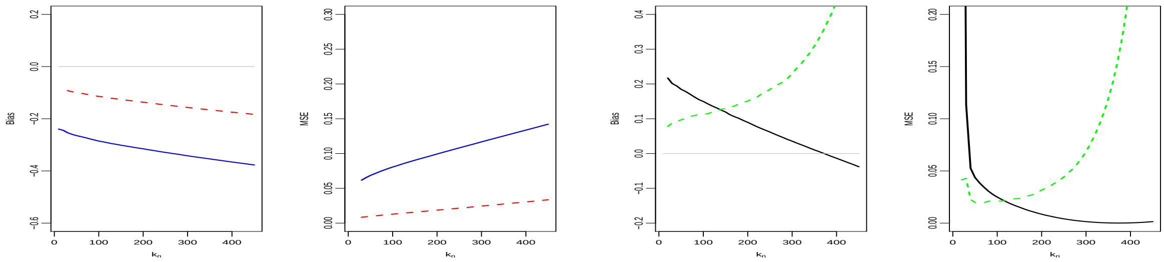
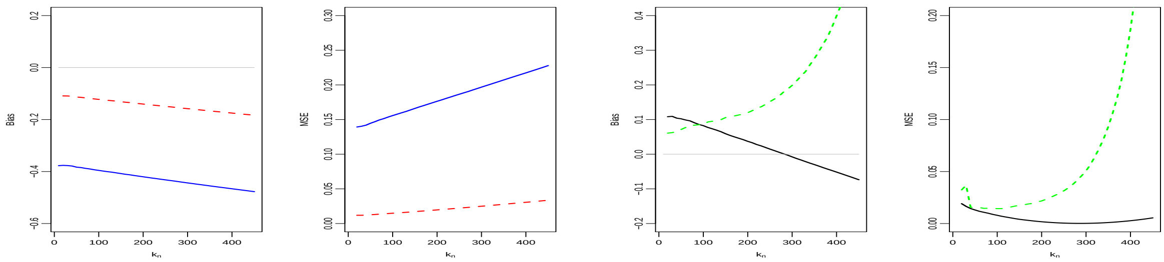


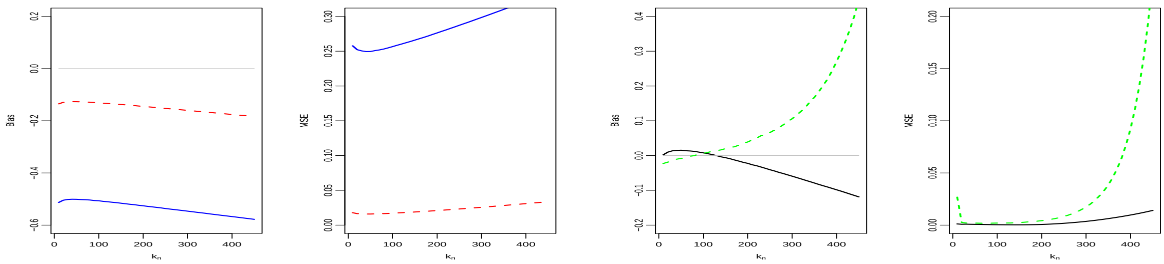
Figure 3: Simulation with X log-Normal censored by C log-Weibull, where $\tau_X = 0.5 < \tau_C = 0.6$ in first line (figures (a)-(b), mild censoring $p = 1$), $\tau_X = 0.5 = \tau_C$ in second line (figures (c)-(d), moderate censoring $p \in]0,1[$), and $\tau_X = 0.5 > \tau_C = 0.4$ in third line (figures (e)-(f), strong censoring $p = 0$). The graphs represent observed bias and MSE of estimators $\hat{\tau}_X$ (blue) and $\hat{\theta}_{X,\hat{\tau}_X}$ (dashed red) in figures (a)-(c)-(e), and relative bias and MSE of estimators $\hat{x}_{p_n}$ (black) and $\hat{x}_{p_n}^{EFG}$ (dashed green) in figures (b)-(d)-(f).



(a) (b)



(c) (d)



(e) (f)

Figure 4: Simulation with X and C in the F model, where $\tau_X = 0.4 < \tau_C = 0.6$ in first line (figures (a)-(b), mild censoring $p = 1$), $\tau_X = 0.5 = \tau_C$ in second line (figures (c)-(d), moderate censoring $p \in]0, 1[$), and $\tau_X = 0.6 > \tau_C = 0.4$ in third line (figures (e)-(f), strong censoring $p = 0$). The graphs represent observed bias and MSE of estimators $\hat{\tau}_X$ (blue) and $\hat{\theta}_{X,\hat{\tau}_X}$ (dashed red) in figures (a)-(c)-(e), and relative bias and MSE of estimators $\hat{x}_{p_n}$ (black) and $\hat{x}_{p_n}^{EFG}$ (dashed green) in figures (b)-(d)-(f).

quite stable zone where k_n is between 55 and 80, we can estimate the probability of non-censoring in the tail by 0.46.

We plot, on the left part of Figure 6, the values of our estimators $\hat{\tau}_X$ (thick blue) and $\hat{\theta}_{X,\hat{\tau}_X}$ (thin red) against k_n. The estimate curve, as a function of k_n, is particularly stable for τ_X (an estimate of 0.12, possibly suggesting a Weibull-tail underlying distribution), but not very stable for θ_X (an estimation between 0.85 and 0.9 if we consider the range of k_n cited above).

On the right part of Figure 6, we plot the values of our estimator $\hat{x}_{p_n}$ (thick blue) of the extreme quantile $\hat{x}_{p_n}$ for the value $p_n = 0.001$, as well as the estimator $\hat{x}_{p_n}^{EFG}$ (thin green) against k_n. We observe that it is quite hard to propose as estimation of x_{p_n} relying on the estimator $\hat{x}_{p_n}^{EFG}$, as it is particularly unstable in the tail. Concerning $\hat{x}_{p_n}$, the choice of the sample fraction is delicate. However, if we consider the quite stable area where k_n is between 55 and 80, as for the estimation of the parameters, then a possible estimation of the extreme quantile is around 28 years.

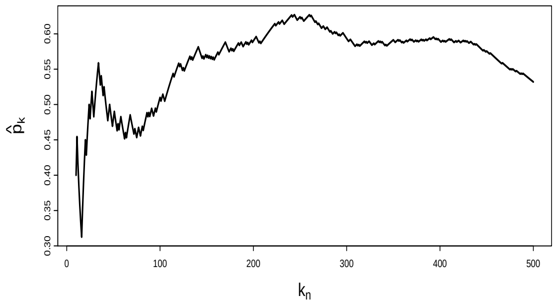


Figure 5: Plot of the proportion of non-censuring $\hat{p}_k$ as a function of k_n.

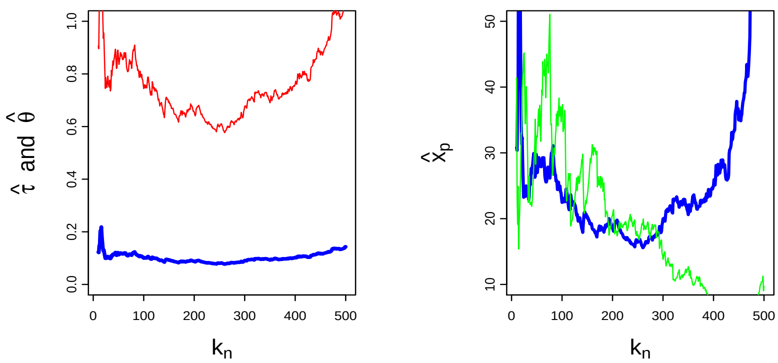


Figure 6: Left: Our estimators $\hat{\tau}_X$ of τ_X (in thick blue) and $\hat{\theta}_{X,\hat{\tau}_X}$ of θ_X (in thin red). Right : Estimators $\hat{x}_{p_n}$ (thick blue) and $\hat{x}_{p_n}^{EFG}$ (thin green) of the extreme quantile x_{p_n}, with $p_n = 0.001$.

A concluding remark could be that estimating extreme quantiles under strong censoring is still a research subject in progress and every new contribution is welcome.

7. Conclusion

In this paper we propose a solution for dealing with tail and extreme quantile estimation of data which are randomly right censored, within a rather large family of distributions encompassing power tail distributions, Weibull-tail distributions, and intermediary situations such as (for instance) log-normal distributions. This family was first introduced in a complete data context in Gardes et al. (2011). Our asymptotic normality results support all possible amounts of censoring in the tail, even very strong ones where the ultimate probability of being censored in the tail is equal to one.

The main two contributions of this work are that very diverse combinations of tails of the censored and censoring distributions are dealt with (not just a combination of tails from the same category), and that tail estimation of log-Weibull-type distributions (not heavier than Pareto tails though) are dealt with as well. The fact that one can estimate the tail parameters of this flexible model, and not just the extreme quantiles, means that the user may consider estimating more elaborated parameters than the extreme quantiles (for instance, expected tail losses $\mathbb{E}(X|X > F_X^-(1-p))$ for small p, of course with additional efforts in order to formally prove convergence results).

Concerning the performances, the bias of our estimators of θ and τ remains a problem, as soon as one moves away from the pure log-Weibull situation. However our opinion is that this bias problem was already present for the original estimators of τ and θ (which inspired ours) in the non-censoring context. This topic of bias reduction still needs to be explored for this family of distributions, even in the non-censored situation. In this paper, we did not try to detail the asymptotic bias, mainly because of the great diversity of situations that our model handled, which already made the exposition a bit complicated. This would require further work.

Finally, a continuation of this work could be to look for estimators of τ and θ which are weighted modifications of their non-censored versions (the estimators in equation (12)), but with varying weights, not the constant weights $D_{k,0}$ and $D_{k,\hat{\tau}_X}$, with in mind a possible improvement in terms of bias and mean-squared error.

References

C. Albert, A. Dutfoy, L. Gardes and S. Girard . An extreme quantile estimator for the log-generalized Weibull-tail model. In *Econometrics and Statistics* **13**, pages 137-174 (2020)

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

J. Beirlant, Y. Goegebeur, J. Segers and J. Teugels . Statistics of extremes: Theory and applications Wiley (2004)

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

J. Beirlant, A. Bardoutsos, T. de Wet and I. Gijbels . Bias reduced tail estimation for censored Pareto type distributions. In *Stat. Prob. Letters* **109**, pages 78-88 (2016)

J. Beirlant, G. Maribe and A. Verster . Penalized bias reduction in extreme value estimation for censored Pareto-type data, and long-tailed insurance applications . In *Insurance : Mathematics and Economics* **78**, pages 114-122 (2018)

J. Beirlant, J. Worms and R. Worms. Asymptotic distribution for an extreme value index estimator in a censorship framework. In *Journal of Statistical Planning and Inference* **202**, pages 31-56 (2019).

M. Bladt, H. Albrecher and J. Beirlant. Combined tail estimation using censored data and expert information. In *Scandinavian Actuarial Journal*, pages 503-525 (2020).

B. Brahimi, D. Meraghni and A. Necir Approximations to the tail index estimator of a heavy-tailed distribution under random censoring and application. In *Mathematical Methods in Statistics* **24**, pages 266-279 (2015)

B. Brahimi, D. Meraghni and A. Necir . Nelson-Aalen tail product-limit process and extreme value index estimation under random censorship. Unpublished manuscript, available on the ArXiv archive : https://arxiv.org/abs/1502.03955v2 (2016)

B. Brahimi, D. Meraghni, A. Necir and L. Soltane. Tail empirical process and a weighted extreme value index estimator for randomly right-censored data. Unpublished manuscript, available on the ArXiv archive : https://arxiv.org/abs/1801.00572 (2018)

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

J. Einmahl, A. Fils-Villetard and A. Guillou . Statistics of extremes under random censoring. In *Bernoulli* **14**, pages 207-227 (2008)

J. El Methni, L. Gardes, S. Girard and A. Guillou . Estimation of extreme quantiles from heavy and light tailed distributions. In *J. Stat. Planning and Inference* **142**, pages 2735-2747 (2012)

L. Gardes, S. Girard and A. Guillou . Weibull tail-distributions revisited: A new look at some tail estimators. In *Journal of Statistical Planning and Inference* **141**, pages 429-444 (2011)

Y. Goegebeur, A. Guillou and J. Qin. Bias-corrected estimation for conditional Pareto-type distributions with random censoring. In *Extremes*, **22**, pages 459-498 (2019 a).

Y. Goegebeur, A. Guillou, and J. Qin . Robust estimation of the Pickands dependence function under random right censoring. In *Insurance: Mathematics and Economics* **87** pages 101-114 (2019 b).

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

L. de Haan and A. Ferreira . Extreme Value Theory : an Introduction. *Springer Science + Business Media* (2006)

E. Hashorva, C. Ling and Z. Peng. Modeling of censored bivariate extremal events. In *Journal of the Korean Statistical Society* **43** pages 323-338 (2014).

L. Hua and H. Joe. Second order regular variation and conditional tail expectation of multiple risks In *Insurance: Mathematics and Economics* **49**, pages 537-546 (2011)

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

P. Ndao., A. Diop, and J-F. Dupuy Nonparametric estimation of the conditional extreme-value index with random covariates and censoring. In *Journal of Statistical Planning and Inference,* **168**, pages 20-37 (2016)

A. Sayah, D. Yahia, and B. Brahimi. On robust tail index estimation under random censorship. In *Afrika Statistika* **9** pages 671-683 (2014).

G. Stupfler . Estimating the conditional extreme-value index in presence of random right-censoring. In *Journal of Multivariate Analysis* **144**, pages 1-24 (2016)

G. Stupfler . On the study of extremes with dependent random right censoring. In *Extremes* **22** pages 97-129 (2019).

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

J. Worms and R. Worms. Moment estimators of the extreme value index for randomly censored data in the Weibull domain of attraction. Unpublished manuscript, available on the ArXiv archive, arXiv:1506.03765 (2015).

J. Worms and R. Worms. Extreme value statistics for censored data with heavy tails under competing risks. In *Metrika* **81 (7)**, pages 849-889 (2018).

J. Worms and R. Worms. Estimation of extremes for Weibull-tail distributions in the presence of random censoring. In *Extremes* **22**, pages 667-704 (2019)

J. Worms and R. Worms. Supplementary Material for : Estimation of extremes for heavy-tailed and light-tailed distributions in the presence of random censoring. Available on the HAL open Archive `https://hal.archives-ouvertes.fr/hal-03046969` (2021)

Appendix

Let us first summarize the contents of the Appendix. It is composed of 5 main parts.

Part A is devoted to the proof of Theorem 1.

Part B is devoted to the proof of Theorem 2.

Part C is devoted to the proof of Theorem 3

Part D is devoted to the proof of Theorem 4.

Part E contains different technical aspects. In particular, the important Lemma 1 and Lemma 2.

The Supplementary Material file contains the proofs of all the Lemmas, and of Propositions 1, 2 and 3.

Recall that L_{nk} is the notation for $\log(n/k)$. Let us introduce the following notations :

$$\Lambda_k = \Lambda_F(Z_{n-k,n}) \quad \text{and} \quad \hat{\Lambda}_k = \hat{\Lambda}_{nX}(Z_{n-k,n}).$$

Appendix A. Proof of Theorem 1

This section details how the asymptotic normality of $\hat{\theta}_{X,\tau_X}$ stems from the combination of properties of the Hill estimator $H_{k,n}$ (relations (A.1), (A.2) and (A.4) below) and of the proportion $\hat{p}_k$ of uncensored data in the tail (Proposition 2 stated next page), via the important decomposition (A.6). Some details are postponed to other sections, in particular the crucial technical Lemma 2 (stated in Appendix E.1) which states the second order properties of the function $p(x) = \mathbb{P}(\delta = 1|Z = x)$. The behavior of the (numerous) remainder terms is detailed in Proposition 3 below.

First, recall that $\hat{\theta}_{X,\tau_X} = \dfrac{H_{k,n}}{D_{k,\tau_X}}$, with

$$H_{k,n} = \frac{1}{k}\sum_{j=1}^{k} \log(Z_{n-j+1,n}) - \log(Z_{n-k,n}) \text{ and } D_{k,\tau_X} = \frac{1}{k}\sum_{j=1}^{k} K_{\tau_X}(\hat{\Lambda}_{nX}(Z_{n-j+1,n})) - K_{\tau_X}(\hat{\Lambda}_{nX}(Z_{n-k,n})).$$

According to Proposition 1, we have $Z_i = H_Z^{-}(\exp(K_{\tau_Z}(E_i)))$, where $E_1, \ldots, E_n$ are n independent standard exponential random variables and (see relation (2)) $H_Z^{-}(x) = x^{\theta_Z}\tilde{l}(x)$, $\tilde{l}$ being RV_0. Hence

$$H_{k,n} = \theta_Z M_n + R_{n,\tilde{l}} \tag{A.1}$$

where

$$M_n := \frac{1}{k}\sum_{j=1}^{k} K_{\tau_Z}(E_{n-i+1,n}) - K_{\tau_Z}(E_{n-k,n}) \quad \text{and} \quad R_{n,\tilde{l}} := \frac{1}{k}\sum_{j=1}^{k} \log\left(\frac{\tilde{l}(\exp(K_{\tau_Z}(E_{n-j+1,n})))}{\tilde{l}(\exp(K_{\tau_Z}(E_{n-k,n})))}\right).$$

By the Renyi representation, we have $E_{n-j+1,n} - E_{n-k} \overset{d}{=} F_{k-j+1,k}$, where $F_1, \ldots, F_k$ are k independent standard exponential random variables. As was done in Gardes et al. (2011) (and borrowing their notations), we have

$$M_n \overset{d}{=} \theta_{n,1}(E_{n-k}) \quad \text{where} \quad \theta_{n,1}(t) := \frac{1}{k}\sum_{j=1}^{k} K_{\tau_Z}(F_i + t) - K_{\tau_Z}(t). \tag{A.2}$$

Introducing, for $q \in \mathbb{N}^*$, (see Lemma 2 of Gardes et al. (2011))

$$\mu_{q,\tau_Z}(t) := \mathbb{E}(\theta_{n,q}(t)) = \int_0^\infty (K_{\tau_Z}(x+t) - K_{\tau_Z}(t))^q \, e^{-x}\, dx = (q!)\, t^{q(\tau_Z-1)}(1 + o(1)) \quad (\text{as } t \to +\infty) \tag{A.3}$$

and $\sigma^2_{1,\tau_Z}(t) := \mu_{2,\tau_Z}(t) - \mu^2_{1,\tau_Z}(t)$, it is proved in Lemma 5 of Gardes et al. (2011) that

$$\sqrt{k}A_{1,n} \xrightarrow{d} N(0,1) \quad \text{where} \quad A_{1,n} := \frac{\theta_{n,1}(E_{n-k}) - \mu_{1,\tau_Z}(E_{n-k})}{\sigma_{1,\tau_Z}(E_{n-k})}. \tag{A.4}$$

Moreover, we prove in Lemma 3 (stated in Appendix E.2) via Taylor's formula that

$$D_{k,\tau_X} = \hat{\Lambda}_k^{\tau_X-1}\hat{p}_k + R_{1,n} \tag{A.5}$$

where $\hat{p}_k$ denotes the proportion of uncensored data among the k upper data values (see Lemma 3 for the definition of the remainder term $R_{1,n}$). Formulas $(A.1)$ and $(A.5)$ thus easily entail the following important intermediate relation :

$$\hat{\theta}_{X,\tau_X} - \theta_X \stackrel{d}{=} \frac{\theta_Z M_n - \theta_X \Lambda_k^{\tau_X - 1}\hat{p}_k}{D_{k,\tau_X}} + \sum_{i=1}^{3} T_{i,n},$$

where

$$\begin{aligned} T_{1,n} &:= \frac{R_{n,\tilde{l}}}{D_{k,\tau_X}} \\ T_{2,n} &:= -\theta_X \frac{R_{1,n}}{D_{k,\tau_X}} \\ T_{3,n} &:= -\theta_X \frac{\hat{\Lambda}_k^{\tau_X-1} - \Lambda_k^{\tau_X-1}}{D_{k,\tau_X}} \hat{p}_k. \end{aligned}$$

Concerning now $\hat{p}_k$, recalling that $a := \theta_Z/\theta_X$, we prove in Lemma 5 (stated in Appendix E.2) that, when $\tau_X \geqslant 0$ and $\tau_C > 0$,

$$\Lambda_k^{\tau_X-1}\hat{p}_k = \left(\frac{a\tau_X}{\tau_Z}\right)^{1-1/\tau_X} E_{n-k}^{\tau_Z(1-1/\tau_X)}\hat{p}_k + R_{2,n}$$

(note that the first term is equal to $E_{n-k}^{-1}\hat{p}_k$ when $0 = \tau_X < \tau_C \leqslant 1$, since then $\tau_Z = \tau_X$ and $a = 1$), and when $\tau_X > 0$ and $\tau_C = 0$,

$$\Lambda_k^{\tau_X-1}\hat{p}_k = (a\tau_X)^{1-1/\tau_X} (\log E_{n-k})^{1-1/\tau_X}\, \hat{p}_k + R_{2,n},$$

where the remainder term $R_{2,n}$ is detailed for each case in the statement of Lemma 5.
Consequently, defining $T_{4,n} := -\theta_X \frac{R_{2,n}}{D_{k,\tau_X}}$, we obtain the following decomposition : when $\tau_X \geqslant 0$ and $\tau_C > 0$

$$\hat{\theta}_{X,\tau_X} - \theta_X \stackrel{d}{=} \frac{\sigma_{1,\tau_Z}(E_{n-k})}{D_{k,\tau_X}} \left(\theta_Z A_{1,n} - \theta_X \frac{\mu_{1,\tau_Z}(E_{n-k})}{\sigma_{1,\tau_Z}(E_{n-k})} \left(\left(\frac{a\tau_X}{\tau_Z}\right)^{1-1/\tau_X} \frac{E_{n-k}^{\tau_Z(1-1/\tau_X)}}{\mu_{1,\tau_Z}(E_{n-k})}\hat{p}_k - \frac{\theta_Z}{\theta_X}\right)\right) + \sum_{i=1}^{4} T_{i,n},$$

and, when $\tau_X > 0$ and $\tau_C = 0$,

$$\hat{\theta}_{X,\tau_X} - \theta_X \stackrel{d}{=} \frac{\sigma_{1,\tau_Z}(E_{n-k})}{D_{k,\tau_X}} \left(\theta_Z A_{1,n} - \theta_X \frac{\mu_{1,\tau_Z}(E_{n-k})}{\sigma_{1,\tau_Z}(E_{n-k})} \left((a\tau_X)^{1-1/\tau_X} \frac{(\log E_{n-k})^{1-1/\tau_X}}{\mu_{1,\tau_Z}(E_{n-k})}\hat{p}_k - \frac{\theta_Z}{\theta_X}\right)\right) + \sum_{i=1}^{4} T_{i,n}.$$

Then, recalling that $\mu_{1,\tau_Z}(t) \sim t^{\tau_Z - 1}$ as $t \to \infty$, we define the following remainder term as (note again that $a\tau_X/\tau_Z = 1$ and $\tau_Z(1 - 1/\tau_X) = -1$ when $\tau_X = 0 < \tau_C$)

$$R_{3,n} := \begin{cases} \left(\frac{a\tau_X}{\tau_Z}\right)^{1-1/\tau_X} \hat{p}_k \left(\frac{(E_{n-k})^{\tau_Z(1-1/\tau_X)}}{\mu_{1,\tau_Z}(E_{n-k})} - L_{nk}^{1-\tau_Z/\tau_X}\right) & \text{when } \tau_X \geqslant 0,\ \tau_C > 0, \\ (a\tau_X)^{1-1/\tau_X}\, \hat{p}_k \left(\frac{(\log E_{n-k})^{1-1/\tau_X}}{\mu_{1,0}(E_{n-k})} - L_{nk}(\log L_{nk})^{1-\tau_Z/\tau_X}\right) & \text{when } 0 = \tau_C < \tau_X. \end{cases}$$

Finally, using the additional fact that, thanks to $(A.3)$, $\frac{\mu_{1,\tau_Z}(E_{n-k})}{\sigma_{1,\tau_Z}(E_{n-k})} \xrightarrow{\mathbb{P}} 1$, we can state the main relation of the proof of Theorem 1 :

$$\hat{\theta}_{X,\tau_X} - \theta_X \stackrel{d}{=} \frac{\mu_{1,\tau_Z}(E_{n-k})}{D_{k,\tau_X}} \left(\theta_Z A_{1,n} - \theta_X A_{2,n}(1 + o_{\mathbb{P}}(1))\right) + \sum_{i=1}^{5} T_{i,n}, \tag{A.6}$$

where the second important term $A_{2,n}$ is defined as

$$A_{2,n} := \begin{cases} \left(\frac{a\tau_X}{\tau_Z}\right)^{1-1/\tau_X} L_{nk}^{1-\tau_Z/\tau_X}\, \hat{p}_k - a & \text{if } \tau_X \geqslant 0 \text{ and } \tau_C > 0, \\ (a\tau_X)^{1-1/\tau_X} L_{nk}(\log L_{nk})^{1-\tau_Z/\tau_X}\, \hat{p}_k - a & \text{if } 0 = \tau_C < \tau_X, \end{cases}$$

and the last remainder term to be introduced is $T_{5,n} := \theta_Z R_{3,n}(1 + o_{\mathbb{P}}(1))$.

We deal with the asymptotic normality of $A_{2,n}$ and the reminder terms $T_{i,n}$ in the following two propositions. Recall that the rate v_n is defined as

$$v_n := \begin{cases} 1 & \text{if } 0 < \tau_X < \tau_C \leqslant 1 \text{ or } 0 < \tau_X = \tau_C < 1 \text{ or } 0 = \tau_X < \tau_C < 1 \\ L_{nk}^{\frac{1}{2}(\frac{\tau_C}{\tau_X}-1)} & \text{if } 0 < \tau_C < \tau_X \leqslant 1 \\ L_{nk}^{-1/2}(\log L_{nk})^{\frac{1}{2}(\frac{1}{\tau_X}-1)} & \text{if } 0 = \tau_C < \tau_X < 1 \end{cases}$$

Proposition 2. *Under the conditions of Theorem 1,*

$$\begin{array}{ll} \textit{if} \quad 0 \leqslant \tau_X < \tau_C \leqslant 1, & \sqrt{k} v_n A_{2,n} = \sqrt{k} v_n (\hat{p}_k - a) = \sqrt{k}(\hat{p}_k - 1) \xrightarrow{\mathbb{P}} 0, \\ \textit{if} \quad 0 < \tau_C < \tau_X \leqslant 1, & \sqrt{k} v_n A_{2,n} = \sqrt{k} v_n \left(\left(\frac{a\tau_X}{\tau_Z}\right)^{1-\frac{1}{\tau_X}} L_{nk}^{1-\tau_Z/\tau_X} \hat{p}_k - a \right) \xrightarrow{d} N\left(0, a^{2-1/\tau_X}\left(\frac{\tau_X}{\tau_C}\right)^{1-1/\tau_X}\right), \\ \textit{if} \quad 0 < \tau_X = \tau_C < 1, & \sqrt{k} v_n A_{2,n} = \sqrt{k}(a^{1-1/\tau_X}\hat{p}_k - a) \xrightarrow{d} N\left(0, a^{2-1/\tau_X}(1 - a^{1/\tau_X})\right), \\ \textit{if} \quad 0 = \tau_C < \tau_X < 1, & \sqrt{k} v_n A_{2,n} = \sqrt{k} v_n \left((a\tau_X)^{1-\frac{1}{\tau_X}} L_{nk} (\log L_{nk})^{1-\frac{1}{\tau_X}} \hat{p}_k - a \right) \xrightarrow{d} N\left(0, a^{2-1/\tau_X}\tau_X^{1-1/\tau_X}\right). \end{array}$$

Proposition 3. *Under the conditions of Theorem 1, for all* $1 \leqslant i \leqslant 5$, $\sqrt{k} v_n T_{i,n} \xrightarrow{\mathbb{P}} 0$*, as* n *tends to infinity.*

The following result is a corollary of Proposition 2 and part of Proposition 3 (concerning the term $T_{2,n}$). As explained at the end of Section 4, the statement of this corollary is helpful for understanding how consistency of an estimator of θ_Z transfers to consistency of our estimator of θ_X.

Corollary 1. *Under the conditions of Theorem 1, we have* $\dfrac{D_{k,\tau_X}}{\mu_{1,\tau_Z}(E_{n-k})} \xrightarrow{\mathbb{P}} a$*, as* n *tends to infinity.*

Indeed, according to $(A.5)$, and since $\mu_{1,\tau_Z}(t) \sim t^{\tau_Z - 1}$ as $t \to \infty$ (see relation (A.3)),

$$\frac{D_{k,\tau_X}}{\mu_{1,\tau_Z}(E_{n-k})} = L_{nk}^{1-\tau_Z} \Lambda_k^{\tau_X - 1} \hat{p}_k (1 + o(1)) \stackrel{d}{=} (A_{2,n} + a)(1 + o(1)) \xrightarrow{\mathbb{P}} a.$$

Of course, Corollary 1 certainly holds with weaker conditions than those of Theorem 1.

Let us end this proof by explaining how the combination of relations (A.6) and (A.4), Propositions 2 and 3, as well as Corollary 1 imply that $\sqrt{k} v_n (\hat{\theta}_{X,\tau_X} - \theta_X) \xrightarrow{d} N(0, v)$ where $v = \theta_X^2 \sigma^2$.

- When $0 \leqslant \tau_X < \tau_C \leqslant 1$, Proposition 2 states that $\sqrt{k} A_{2,n}$ converges to 0. Hence, the leading term in $(A.6)$ is $\sqrt{k} A_{1,n}$ which converges in distribution to $N(0, 1)$ (see $(A.4)$), and we thus obtain the desired value of $v = (\frac{1}{a})^2 \theta_Z^2 = \theta_X^2$.

- When $0 < \tau_X = \tau_C < 1$, Proposition 2 states that $\sqrt{k} A_{2,n} \xrightarrow{d} N\left(0, a^{2-1/\tau_X}(1 - a^{1/\tau_X})\right)$. Moreover $\sqrt{k} A_{1,n}$ converges in distribution to $N(0, 1)$. Since $A_{1,n}$ and $A_{2,n}$ are independent (under our independent censoring setting), we obtain as desired

$$v = \frac{\theta_Z^2}{a^2} + \frac{\theta_X^2}{a^2} a^{2-1/\tau_X}(1 - a^{1/\tau_X}) = \theta_X^2 + \theta_X^2(a^{-1/\tau_X} - 1) = \theta_X^2 a^{-1/\tau_X}.$$

- In the other two cases, since $v_n \to 0$, $\sqrt{k} v_n A_{1,n}$ converges in probability to 0, and on the other hand Proposition 2 states that $\sqrt{k} v_n A_{2,n}$ converges in distribution to $N(0, D)$ with a variance described above, and it is not difficult to check that $(\frac{1}{a})^2 \theta_X^2 D$ equals to $\theta_X^2 \sigma^2$ as stated.

Appendix B. Proof of Theorem 2

The proof is very similar to the previous one. First, recall that $\hat{\tau}_X = \frac{HH_{k,n}}{D_{k,0}}$. Concerning the numerator, we have by Proposition 1 that $Z_i = H_Z^{-}(\exp(K_{\tau_Z}(E_i)))$, where $E_1, \ldots, E_n$ are standard exponential, and thus

$$HH_{k,n} := \frac{1}{k_n} \sum_{j=1}^{k_n} \log\log(Z_{n-j+1,n}) - \log\log(Z_{n-k_n,n}) = LM_n + RR_{n,\tilde{l}} \tag{B.1}$$

where

$$LM_n := \frac{1}{k} \sum_{j=1}^{k} \log(K_{\tau_Z}(E_{n-i+1,n})) - \log(K_{\tau_Z}(E_{n-k,n})) \quad \text{and} \quad RR_{n,\tilde{l}} := \frac{1}{k} \sum_{j=1}^{k} \log \left(\frac{1 + \frac{\log(\tilde{l}(\exp(K_{\tau_Z}(E_{n-j+1,n}))))}{\theta_Z K_{\tau_Z}(E_{n-j+1,n})}}{1 + \frac{\log(\tilde{l}(\exp(K_{\tau_Z}(E_{n-k,n}))))}{\theta_Z K_{\tau_Z}(E_{n-k,n})}} \right).$$

By the Renyi representation, for some independent standard exponential random variables $F_1, \ldots, F_k$ we have

$$LM_n \stackrel{d}{=} l\theta_{n,1}(E_{n-k}) \quad \text{where} \quad l\theta_{n,1}(t) := \frac{1}{k}\sum_{j=1}^{k} \log(K_{\tau_Z}(F_i + t)) - \log(K_{\tau_Z}(t)). \tag{B.2}$$

Introducing, for $q \in \mathbb{N}^*$,

$$l\mu_{q,\tau_Z}(t) := \mathbb{E}(l\theta_{n,q}(t)) = \int_0^\infty (\log(K_{\tau_Z}(x+t)) - \log(K_{\tau_Z}(t)))^q \, e^{-x} \, dx$$

and $l\sigma^2_{1,\tau_Z}(t) := l\mu_{2,\tau_Z}(t) - l\mu^2_{1,\tau_Z}(t)$, we have

$$l\mu_{q,\tau_Z}(t) = \begin{cases} (q!)\tau_Z^q t^{-q}(1+o(1)) & \text{if } \tau_Z \neq 0, \\ (q!)t^{-q}(\log(t))^{-q}(1+o(1)) & \text{if } \tau_Z = 0. \end{cases} \tag{B.3}$$

We can then prove that (the proof is similar to that of Lemma 5 in Gardes et al. (2011))

$$\sqrt{k} LA_{1,n} \xrightarrow{d} N(0,1) \quad \text{where} \quad LA_{1,n} := \frac{l\theta_{n,1}(E_{n-k}) - l\mu_{1,\tau_Z}(E_{n-k})}{l\sigma_{1,\tau_Z}(E_{n-k})}. \tag{B.4}$$

Concerning now the denominator, we prove in Lemma 3 (stated in Appendix E.2) that

$$D_{k,0} := \frac{1}{k_n}\sum_{j=1}^{k_n} \log(\hat{\Lambda}_{nX}(Z_{n-j+1,n})) - \log(\hat{\Lambda}_{nX}(Z_{n-k_n,n})) = \hat{\Lambda}_k^{-1}\hat{p}_k + R_{1,n}, \tag{B.5}$$

where

$$R_{1,n} = \frac{1}{k}\sum_{j=1}^{k}\left(\log\left(1 + \frac{\hat{\Delta}_{j,k}}{\hat{\Lambda}_k}\right) - \frac{\hat{\Delta}_{j,k}}{\hat{\Lambda}_k}\right)$$

where the $\hat{\Delta}_{j,k}$ are defined in Lemma 3 and $\hat{p}_k$ denotes the proportion of uncensored data in the tail. From now on we consider that $\tau_X \neq 0$ (see Remark 8 below for the $\tau_X = 0$ case). Formulas $(B.1)$ and $(B.5)$ easily entail the following important intermediary relation :

$$\hat{\tau}_X - \tau_X \stackrel{d}{=} \frac{LM_n - \tau_X \Lambda_k^{-1}\hat{p}_k}{D_{k,0}} + \sum_{i=1}^{3} TT_{i,n},$$

where

$$\begin{aligned} TT_{1,n} &:= \frac{RR_{n,\tilde{l}}}{D_{k,0}} \\ TT_{2,n} &:= -\tau_X \frac{R_{1,n}}{D_{k,0}} \\ TT_{3,n} &:= -\tau_X(\hat{\Lambda}_k^{-1} - \Lambda_k^{-1})(D_{k,0})^{-1}\hat{p}_k. \end{aligned}$$

Moreover, we prove in Lemma 6 (stated in Appendix E.2) that, when $\tau_X > 0$ and $\tau_C > 0$ (the case $\tau_X > 0$ and $\tau_C = 0$ is omitted for brevity),

$$\Lambda_k^{-1}\hat{p}_k = \left(\frac{a\tau_X}{\tau_Z}\right)^{-1/\tau_X} E_{n-k}^{-\tau_Z/\tau_X}\hat{p}_k + RR_{2,n},$$

the expression for the remainder term $RR_{2,n}$ being detailed for each case in the statement of Lemma 6.

Consequently, defining $TT_{4,n} := -\tau_X \frac{RR_{2,n}}{D_{k,0}}$, we obtain the following decomposition : when $\tau_X > 0$ and $\tau_C > 0$

$$\hat{\tau}_X - \tau_X \stackrel{d}{=} \frac{l\sigma_{1,\tau_Z}(E_{n-k})}{D_{k,0}}\left(LA_{1,n} - \tau_X \frac{l\mu_{1,\tau_Z}(E_{n-k})}{l\sigma_{1,\tau_Z}(E_{n-k})}\left(\left(\frac{a\tau_X}{\tau_Z}\right)^{-1/\tau_X} \frac{E_{n-k}^{-\tau_Z/\tau_X}}{l\mu_{1,\tau_Z}(E_{n-k})}\hat{p}_k - \frac{1}{\tau_X}\right)\right) + \sum_{i=1}^{4} T_{i,n}.$$

But $l\mu_{1,\tau_Z}(t) \sim \tau_Z t^{-1}$, so we define the following remainder term as

$$RR_{3,n} := \left(\frac{a\tau_X}{\tau_Z}\right)^{-1/\tau_X} \hat{p}_k \left(\frac{(E_{n-k})^{-\tau_Z/\tau_X}}{l\mu_{1,\tau_Z}(E_{n-k})} - \frac{1}{\tau_Z} L_{nk}^{1-\tau_Z/\tau_X}\right).$$

Finally, using the additional fact that $\frac{l\mu_{1,\tau_Z}(E_{n-k})}{l\sigma_{1,\tau_Z}(E_{n-k})} \xrightarrow{\mathbb{P}} 1$, we can state the main relation of the proof of Theorem 2 :

$$\hat{\tau}_X - \tau_X \stackrel{d}{=} \frac{l\mu_{1,\tau_Z}(E_{n-k})}{D_{k,0}} \left(LA_{1,n} - a^{-1}A_{2,n}(1 + o_{\mathbb{P}}(1))\right) + \sum_{i=1}^{5} TT_{i,n}, \tag{B.6}$$

where $LA_{1,n}$ is defined in (B.4), the second main term $A_{2,n}$ is defined in section Appendix A and the last remainder term to be introduced is $TT_{5,n} := -\tau_X RR_{3,n}(1 + o_{\mathbb{P}}(1))$. The asymptotic normality of $A_{2,n}$ is dealt with in Proposition 2. Concerning the remainder terms $TT_{i,n}$, we prove the following proposition :

Proposition 4. *Under the conditions of Theorem 1, for all* $1 \leqslant i \leqslant 5$, $\sqrt{k}v_n TT_{i,n} \xrightarrow{\mathbb{P}} 0$, *as* n *tends to infinity.*

The proof of Proposition 4 is very similar to the proof of Proposition 3.

The following statement is a Corollary of Propositions 2 and 4, in the same way that Corollary 1 was deduced from Propositions 2 and 3.

Corollary 2. *Under the conditions of Theorem 1, when* $\tau_X \neq 0$ *we have*

$$\frac{D_{k,0}}{l\mu_{1,\tau_Z}(E_{n-k})} \xrightarrow{\mathbb{P}} \frac{1}{\tau_X}$$

and, when $0 = \tau_X < \tau_C$, *we have as* $n \to \infty$

$$\frac{D_{k,0}}{l\mu_{1,0}(E_{n-k})} = (\log L_{nk})(1 + o_{\mathbb{P}}(1)).$$

The proof of Theorem 2 can be concluded in the same way as was that of Theorem 1. Details are omitted. □

Remark 8. *In the case* $0 = \tau_X < \tau_C$, *we have* $D_{k,0}/l\mu_{1,0}(E_{n-k,n}) \stackrel{\mathbb{P}}{\sim} \log L_{nk}$, *and thus the estimator* $\hat{\tau}_X \stackrel{d}{=} l\theta_{n,1}(E_{n-k}) / D_{k,0} + TT_{1,n}$ *is contiguous to* $l\mu_{1,0}(E_{n-k}) / D_{k,0} + TT_{1,n}$, *which is itself equivalent in probability to* $1/\log L_{nk}$. *Thus only the consistency and rate of convergence of* $\hat{\tau}_X$ *is obtained in this case.*

Appendix C. Proof of Theorem 3

Recall that $\hat{\theta}_{X,\hat{\tau}_X} = H_{k,n}/D_{k,\hat{\tau}_X}$ where

$$H_{k,n} = \frac{1}{k}\sum_{j=1}^{k} \log(Z_{n-j+1,n}) - \log(Z_{n-k,n}) \text{ and } D_{k,\hat{\tau}_X} = \frac{1}{k}\sum_{j=1}^{k} K_{\hat{\tau}_X}(\hat{\Lambda}_{nX}(Z_{n-j+1,n})) - K_{\hat{\tau}_X}(\hat{\Lambda}_{nX}(Z_{n-k,n})).$$

Moreover

$$\log\left(\frac{\hat{\theta}_{X,\hat{\tau}_X}}{\theta_X}\right) = \log\left(\frac{\hat{\theta}_{X,\hat{\tau}_X}}{\hat{\theta}_{X,\tau_X}}\right) + \log\left(\frac{\hat{\theta}_{X,\tau_X}}{\theta_X}\right). \tag{C.1}$$

Theorem 1 and the delta-method yields that the second term of the right-hand side in $(C.1)$ satisfies

$$\sqrt{k}v_n \log\left(\frac{\hat{\theta}_{X,\tau_X}}{\theta_X}\right) \xrightarrow{d} N\left(0, \sigma^2\right). \tag{C.2}$$

Now let us treat the first term. Since $D_{k,\tau_X} = (\hat{\Lambda}_k)^{\tau_X - 1}\hat{p}_k + R_{1,n}$ (see Lemma 3) and, similarly, $D_{k,\hat{\tau}_X} = (\hat{\Lambda}_k)^{\hat{\tau}_X - 1}\hat{p}_k + \hat{R}_{1,n}$, where $\hat{R}_{1,n}$ is obtained by replacing τ_X by $\hat{\tau}_X$ in the expression for $R_{1,n}$, we obtain

$$\log\left(\frac{\hat{\theta}_{X,\tau_X}}{\hat{\theta}_{X,\hat{\tau}_X}}\right) = (\hat{\tau}_X - \tau_X)\log(\hat{\Lambda}_k) - \log\left(1 + \frac{R_{1,n}}{\hat{\Lambda}_k^{\tau_X - 1}\hat{p}_k}\right) + \log\left(1 + \frac{\hat{R}_{1,n}}{\hat{\Lambda}_k^{\hat{\tau}_X - 1}\hat{p}_k}\right).$$

Let us study separately the first two terms in the expression above (the third one being similar to the second one). The starting point is

$$(\hat{\tau}_X - \tau_X)\log(\hat{\Lambda}_k) = (\hat{\tau}_X - \tau_X)\log(\Lambda_k) + (\hat{\tau}_X - \tau_X)\log\left(\frac{\hat{\Lambda}_k}{\Lambda_k}\right).$$

Let us continue with the case $\tau_X \neq 0$ and $\tau_C \neq 0$ (the case $0 = \tau_C < \tau_X$ being similar and the case $0 = \tau_X < \tau_C$ being excluded, see Remark 9 below).

Since $\sqrt{k}v_n(\hat{\tau}_X - \tau_X) \xrightarrow{d} N\left(0, \sigma^2\tau_X^2\right)$ (Theorem 2), and, according to Lemma 7, $\log(\Lambda_k) = \frac{\tau_Z}{\tau_X}(\log L_{nk})(1 + o_{\mathbb{P}}(1))$, we obtain that

$$\frac{\sqrt{k}v_n}{\log L_{nk}}(\hat{\tau}_X - \tau_X)\log(\Lambda_k) \xrightarrow{d} N\left(0, \sigma^2\tau_Z^2\right)$$

and $\frac{\sqrt{k}v_n}{\log L_{nk}}(\hat{\tau}_X - \tau_X)\log\left(\frac{\hat{\Lambda}_k}{\Lambda_k}\right) = o_{\mathbb{P}}(1)$ (because $\frac{\hat{\Lambda}_k}{\Lambda_k} = O_{\mathbb{P}}(1)$).

Now, $\log\left(1 + \frac{R_{1,n}}{(\hat{\Lambda}_k)^{\tau_X - 1}\hat{p}_k}\right) = \frac{R_{1,n}}{(\hat{\Lambda}_k)^{\tau_X - 1}\hat{p}_k}(1 + o_{\mathbb{P}}(1))$, and we prove in Proposition 3 that $\sqrt{k}v_n\frac{R_{1,n}}{(\hat{\Lambda}_k)^{\tau_X-1}\hat{p}_k} = o_{\mathbb{P}}(1)$. Hence $\frac{\sqrt{k}v_n}{\log L_{nk}}\log\left(1 + \frac{R_{1,n}}{(\hat{\Lambda}_k)^{\tau_X-1}\hat{p}_k}\right) = o_{\mathbb{P}}(1)$. This ensures that

$$\frac{\sqrt{k}v_n}{\log L_{nk}}\log\left(\frac{\hat{\theta}_{X,\tau_X}}{\hat{\theta}_{X,\hat{\tau}_X}}\right) \xrightarrow{d} N\left(0, \sigma^2\tau_Z^2\right).$$

Finally, $(C.1)$ and $(C.2)$ yield

$$\frac{\sqrt{k}v_n}{\log L_{nk}}\log\left(\frac{\hat{\theta}_{X,\hat{\tau}_X}}{\theta_X}\right) \xrightarrow{d} N\left(0, \sigma^2\tau_Z^2\right).$$

This entails the announced asymptotic normality, via the delta-method. □

Remark 9. *In the case $\tau_X = 0$, $\log(\Lambda_k) = a(\log L_{nk})(1 + o_{\mathbb{P}}(1))$, according to Lemma 7. Hence, $\hat{\tau}_X\log(\Lambda_k)$ does not converge to 0, in this case. This is why $\tau_X = 0$ is excluded from the asymptotic result of $\hat{\theta}_{X,\hat{\tau}_X}$.*

Appendix D. Proof of Theorem 4

Recall that $x_{p_n} = \bar{F}_X^{-}(p_n) = H_X^{-}(\exp(K_{\tau_X}(-\log p_n)))$ and

$$\hat{x}_{p_n} = Z_{n-k,n}\exp\left(\hat{\theta}_{X,\hat{\tau}_X}\left(K_{\hat{\tau}_X}(-\log(p_n)) - K_{\hat{\tau}_X}(\hat{\Lambda}_k)\right)\right)$$

where $H_X^{-}(x) = x^{\theta_X}\bar{l}_X(x)$, and $\bar{l}_X$ is slowly varying at infinity. Moreover, since $Z_{n-k,n} = \bar{F}_X^{-}(\exp(-\Lambda_k))$, it is easy to prove that

$$\begin{aligned}
\log\left(\tfrac{\hat{x}_{p_n}}{x_{p_n}}\right) = \;& \hat{\theta}_{X,\hat{\tau}_X}\left\{(K_{\hat{\tau}_X}(-\log(p_n)) - K_{\hat{\tau}_X}(\Lambda_k)) - (K_{\tau_X}(-\log(p_n)) - K_{\tau_X}(\Lambda_k))\right\} \\
& +(\hat{\theta}_{X,\hat{\tau}_X} - \theta_X)K_{\tau_X}(-\log(p_n)) + \hat{\theta}_{X,\hat{\tau}_X}\left(K_{\hat{\tau}_X}(\Lambda_k) - K_{\hat{\tau}_X}(\hat{\Lambda}_k)\right) \\
& -(\hat{\theta}_{X,\hat{\tau}_X} - \theta_X)K_{\tau_X}(\Lambda_k) + \log\left(\tfrac{\bar{l}_X(\exp(K_{\tau_X}(\Lambda_k)))}{\bar{l}_X(\exp(K_{\tau_X}(-\log(p_n))))}\right) \\
& =: Q_1 + Q_2 + Q_3 + Q_4 + Q_5.
\end{aligned}$$

Let us treat separately these five terms, in the case $\tau_X \neq 0$ and $\tau_C \neq 0$, the case $0 = \tau_X < \tau_C$ being similar. Note that Q_1 will turn out to be the main term.

Recall that

$$L_k := \begin{cases} (a\tau_X/\tau_Z)^{1/\tau_X}(L_{nk})^{\tau_Z/\tau_X} & \text{if } \tau_X \neq 0 \text{ and } \tau_C \neq 0, \\ (a\tau_X)^{1/\tau_X}(\log L_{nk})^{1/\tau_X} & \text{if } \tau_X \neq 0 \text{ and } \tau_C = 0. \end{cases}$$

Consider the temporary notations

$$\sigma_n := \left(\sqrt{k}v_n\right)^{-1} \quad \text{and} \quad w_n := \int_{L_k}^{-\log(p_n)} u^{\tau_x - 1}\log u\, du.$$

By integration by parts, and under assumption (11) (which implies that $L_k = o(-\log(p_n))$), we can prove that

$$w_n = \frac{1}{\tau_X}\log(\log(1/p_n))\left(-\log(p_n)\right)^{\tau_X}(1 + o(1)), \tag{D.1}$$

and similarly $\tilde{w}_n := \int_{L_k}^{-\log(p_n)} u^{\tau_x - 1}\log^2 u\, du = \frac{1}{\tau_X}(\log(\log(1/p_n)))^2\left(-\log(p_n)\right)^{\tau_X}(1 + o(1))$.

- Let us prove that $\sigma_n^{-1}w_n^{-1}Q_1$ converges in distribution to $\mathcal{N}(0,\theta_X^2\tau_X^2\sigma^2)$, which (via (D.1)) will imply that

$$\frac{\sqrt{k}v_n}{\log\log(1/p_n)(-\log p_n)^{\tau_X}}Q_1 \xrightarrow{d} \mathcal{N}(0,\theta_X^2\sigma^2). \tag{D.2}$$

According to Theorem 2, $\hat\tau_X = \tau_X + \sigma_n\xi_n$, where ξ_n converges in distribution to $\mathcal{N}(0,\tau_X^2\sigma^2)$. Hence,

$$\begin{aligned} Q_1 &= \hat\theta_{X,\hat\tau_X}\left(\int_{\Lambda_k}^{-\log p_n} u^{\tau_X+\sigma_n\xi_n-1}du - \int_{\Lambda_k}^{-\log p_n} u^{\tau_X-1}du\right) \\ &= \hat\theta_{X,\hat\tau_X}\left(\int_{L_k}^{-\log p_n} u^{\tau_X-1}(u^{\sigma_n\xi_n}-1)du - \int_{L_k}^{\Lambda_k} u^{\tau_X-1}(u^{\sigma_n\xi_n}-1)du\right). \end{aligned}$$

Let us introduce $\phi(x)=e^x-1-x$. Consequently,

$$Q_1 = \sum_{i=1}^4 Q_1^{(i)},$$

where

$$\begin{aligned} Q_1^{(1)} &= \hat\theta_{X,\hat\tau_X}\int_{L_k}^{-\log p_n} u^{\tau_X-1}\phi(\sigma_n\xi_n\log u)du \\ Q_1^{(2)} &= \hat\theta_{X,\hat\tau_X}\sigma_n\xi_n\int_{L_k}^{-\log p_n} u^{\tau_X-1}\log u\ du \\ Q_1^{(3)} &= -\hat\theta_{X,\hat\tau_X}\int_{L_k}^{\Lambda_k} u^{\tau_X-1}\phi(\sigma_n\xi_n\log u)du \\ Q_1^{(4)} &= -\hat\theta_{X,\hat\tau_X}\sigma_n\xi_n\int_{L_k}^{\Lambda_k} u^{\tau_X-1}\log u\ du \end{aligned}$$

Now, there exists $\eta>0$, such that $x<\log\eta$ implies that $|\phi(x)|<(\eta/2)x^2$. As a consequence, since $\sigma_n\log\log(1/p_n)\to 0$ and $\sigma_n\log L_k\to 0$ (according to (10) and (11)),

$$|Q_1^{(1)}| < \hat\theta_{X,\hat\tau_X}\frac{\eta}{2}\sigma_n^2\xi_n^2\int_{L_k}^{-\log p_n} u^{\tau_X-1}(\log u)^2\ du = \eta\, O_{\mathbb{P}}(1)\,\sigma_n^2\tilde w_n.$$

Hence, via (10) and the previous approximations of w_n and $\tilde w_n$,

$$\sigma_n^{-1}w_n^{-1}|Q_1^{(1)}| < \eta\, O_{\mathbb{P}}(1)\,\sigma_n\tilde w_n/w_n = \eta\, O_{\mathbb{P}}(1)\,\sigma_n\log\log(1/p_n) \xrightarrow{\mathbb{P}} 0.$$

Concerning $Q_1^{(2)}$, we have

$$\sigma_n^{-1}w_n^{-1}Q_1^{(2)} = \hat\theta_{X,\hat\tau_X}\xi_n \xrightarrow{d} \mathcal{N}(0,\theta_X^2\tau_X^2\sigma^2).$$

Let us now consider $Q_1^{(3)}$. We proceed as for $Q_1^{(1)}$ to obtain

$$\begin{aligned} \sigma_n^{-1}w_n^{-1}|Q_1^{(3)}| &< \hat\theta_{X,\hat\tau_X}\tfrac{\eta}{2}\sigma_n\xi_n^2\frac{\int_{L_k}^{\Lambda_k}u^{\tau_X-1}(\log u)^2\ du}{\int_{L_k}^{-\log p_n}u^{\tau_X-1}\log u\ du} \\ &< \hat\theta_{X,\hat\tau_X}\tfrac{\eta}{2}\sigma_n\max(\log\Lambda_k,\log L_k)\xi_n^2\frac{\int_{L_k}^{\Lambda_k}u^{\tau_X-1}\log u\ du}{\int_{L_k}^{-\log p_n}u^{\tau_X-1}\log u\ du}. \end{aligned}$$

Since $\sigma_n\log\Lambda_k \xrightarrow{\mathbb{P}} 0$ (this is an easy consequence of assumption (11) and Lemma 7), the right hand-side tends to 0, according to Lemma 8 and assumption (11).

Concerning $Q_1^{(4)}$, Lemma 8 and assumption (11) entails that $\sigma_n^{-1}w_n^{-1}Q_1^{(4)}$ tends to 0. This completes the proof of (D.2).

- Let us prove that $\sigma_n^{-1}w_n^{-1}Q_2 = o_{\mathbb{P}}(1)$: according to Theorem 3,

$$Q_2 = \sigma_n(\log L_{nk})K_{\tau_X}(-\log(p_n))\delta_n,$$

where δ_n converges to a gaussian distribution. Hence,

$$\sigma_n^{-1}w_n^{-1}Q_2 = \frac{(\log L_{nk})\ K_{\tau_X}(-\log(p_n))}{\int_{L_k}^{-\log(p_n)}u^{\tau_x-1}\log u\ du}\delta_n,$$

and assumption (11) yields the result.

- In order to prove that $\sigma_n^{-1}w_n^{-1}Q_3 = o_{\mathbb{P}}(1)$, we obtain via a Taylor expansion that

$$\sigma_n^{-1}w_n^{-1}|Q_3| = \hat\theta_{X,\hat\tau_X}\sqrt{k}|\Lambda_k-\hat\Lambda_k|\left|K'_{\hat\tau_X}(\hat T_k)\frac{v_n}{w_n}\right|$$

where $\hat{T}_k$ is a value between Λ_k and $\hat{\Lambda}_k$. The fact that $\sqrt{k}|\Lambda_k - \hat{\Lambda}_k| = O_{\mathbb{P}}(1)$ (see Lemma 7 in Worms and Worms (2019)) and assumption (11) yields the result.

- Let us prove that $\sigma_n^{-1} w_n^{-1} Q_4 = o_{\mathbb{P}}(1)$: as above (see treatment of Q_2)

$$Q_4 = \sigma_n \log L_{nk} K_{\tau_X}(\Lambda_k)\delta_n,$$

where δ_n converges to a gaussian distribution. Moreover $K_{\tau_X}(\Lambda_k) \stackrel{d}{=} a K_{\tau_Z}(L_{nk})(1+o_{\mathbb{P}}(1))$ (see Lemma 4 (i)). Hence

$$\sigma_n^{-1} w_n^{-1} Q_4 \stackrel{d}{=} a \frac{K_{\tau_Z}(L_{nk})\ \log L_{nk}}{\int_{L_k}^{-\log(p_n)} u^{\tau_x - 1} \log u \ du}\ \delta_n(1 + o_{\mathbb{P}}(1)).$$

Assumption (11) yields the result.

- Let us finally prove that $\sigma_n^{-1} w_n^{-1} Q_5 = o_{\mathbb{P}}(1)$: recall that

$$\begin{aligned} Q_5 &= \log\left(\tfrac{\bar{l}_X(\exp(K_{\tau_X}(\Lambda_k)))}{\bar{l}_X(\exp(K_{\tau_X}(-\log(p_n))))}\right) \\ &= \log\left(\tfrac{\bar{l}_X(\exp(K_{\tau_X}(L_k)))}{\bar{l}_X(\exp(K_{\tau_X}(-\log(p_n))))}\right) + \log\left(\tfrac{\bar{l}_X(\exp(K_{\tau_X}(\Lambda_k)))}{\bar{l}_X(\exp(K_{\tau_X}(L_k)))}\right) \\ &= Q_5^{(1)} + Q_5^{(2)}. \end{aligned}$$

Concerning $Q_5^{(1)}$, we know that $\bar{l}_X$ satisfies the SR2 condition (see Remark 10). Hence

$$\begin{aligned} -Q_5^{(1)} &= \log\left(\tfrac{\bar{l}_X(\lambda_n x_n)}{\bar{l}_X(x_n)}\right) \\ &= \bar{b}_X(x_n) K_{\theta_X \rho_X}(\lambda_n)(1 + o_{\mathbb{P}}(1)), \end{aligned}$$

where $|\bar{b}_X| \in RV_{\theta_X \rho_X}$, $\lambda_n = \frac{\exp(K_{\tau_X}(-\log(p_n)))}{\exp(K_{\tau_X}(L_k))}$ and $x_n = \exp(K_{\tau_X}(L_k))$. Moreover, since λ_n tends to $+\infty$, as n tends to infinity (because $\frac{K_{\tau_X}(L_k)}{K_{\tau_X}(-\log(p_n))}$ tends to 0 under assumption (11)), we obtain that $K_{\theta_X \rho_X}(\lambda_n)$ tends to $-1/(\theta_X \rho_X)$. Moreover, $\sqrt{k} v_n \bar{b}_X(\exp(K_{\tau_X}(L_k)))$ tends to 0 under the appropriate assumption among $H_2, \ldots, H_5$. Hence,

$$\sqrt{k} v_n \frac{Q_5^{(1)}}{K_{\tau_X}(-\log(p_n))} = \sqrt{k} v_n \bar{b}_X(\exp(K_{\tau_X}(L_k))) \frac{K_{\rho_X}(\lambda_n)}{K_{\tau_X}(-\log(p_n))},$$

tends to 0. Then,

$$\sigma_n^{-1} w_n^{-1} Q_5^{(1)} = \sqrt{k} v_n \frac{Q_5^{(1)}}{K_{\tau_X}(-\log(p_n))} \frac{K_{\tau_X}(-\log p_n)}{\int_{L_k}^{-\log(p_n)} u^{\tau_x - 1} \log u \ du},$$

which tends to 0 thanks to $(D.1)$.

Similarly, we have

$$\begin{aligned} Q_5^{(2)} &= \log\left(\tfrac{\bar{l}_X(\lambda_n x_n)}{\bar{l}_X(x_n)}\right) \\ &= \bar{b}_X(x_n) K_{\theta_X \rho_X}(\lambda_n)(1 + o_{\mathbb{P}}(1)), \end{aligned}$$

where $x_n = \exp(K_{\tau_X}(L_k))$ and

$$\lambda_n = \frac{\exp(K_{\tau_X}(\Lambda_k))}{\exp(K_{\tau_X}(L_k))} = \exp(\tau_X^{-1}(\Lambda_k^{\tau_X} - L_k^{\tau_X})) = \exp(cst.L_{nk}^{\tau_Z - \alpha}(1 + o(1))),$$

where, according to Lemma 4, the constant above is negative and

$$\alpha = \begin{cases} \tau_Z & \text{when either } \tau_X = \tau_C, \text{ or } \tau_X \neq \tau_C \text{ and } r \leqslant 0, \\ \tau_Z(1 - r) & \text{when } \tau_X \neq \tau_C \text{ and } r \in]0, 1[. \end{cases}$$

In the case where $\alpha = \tau_Z$, $K_{\theta_X \rho_X}(\lambda_n)$ converges to a constant. Hence we obtain, for the term $Q_5^{(1)}$, that

$$\sqrt{k} v_n \frac{Q_5^{(2)}}{K_{\tau_X}(-\log(p_n))} \xrightarrow{\mathbb{P}} 0.$$

Therefore $\sigma_n^{-1} w_n^{-1} Q_5^{(2)} \xrightarrow{\mathbb{P}} 0$, thanks to $(D.1)$.

In the case where $\alpha = \tau_Z(1-r)$, we have $K_{\theta_X\rho_X}(\lambda_n) = O(1)\exp(cst.L_{nk}^{r\tau_Z}(1+o(1)))$, where here the constant is positive. Moreover, for some small $\delta > 0$,

$$\bar{b}_X(x_n) = \exp((\theta_X\rho_X + \delta)K_{\tau_X}(L_k))o(1) = \exp\left((\theta_X\rho_X+\delta).cst.L_{nk}^{\tau_Z}(1+o(1))\right)o(1),$$

where the constant above is positive. Consequently, $\sqrt{k}v_n\bar{b}_X(x_n)K_{\theta_X\rho_X}(\lambda_n)$ tends to 0 according to the appropriate assumption among $H_2,\ldots,H_5$. To conclude, we proceed as in the previous case. $\square$

Appendix E. Technical aspects

The Lemmas stated in this section are proved in the Supplementary Material document.

Appendix E.1. Details about second order conditions and censoring probabilities

Recall that

$$\bar{F}_X(x) = \exp(-K_{\tau_X}^{-}(\log(H_X(x)))) \quad \text{and} \quad \bar{F}_C(x) = \exp(-K_{\tau_C}^{-}(\log(H_C(x))))$$

where

$$H_X^{-}(x) = x^{\theta_X}\bar{l}_X(x)\ ,\ H_C^{-}(x) = x^{\theta_C}\bar{l}_C(x)\ ,\ H_X(x) = x^{1/\theta_X}l_X(x)\ ,\ H_C(x) = x^{1/\theta_C}l_C(x).$$

Moreover (see Proposition 1),

$$\mathbb{P}(Z > x) = \exp(-K_{\tau_Z}^{-}(\log(H_Z(x)))),$$

where $H_Z^{-}(x) = x^{\theta_Z}\tilde{l}(x)$ and $\tilde{l}$ is slowly varying. This implies that $H_X \circ H_Z^{-}(x) = x^a l(x)$, with l a slowly varying function and $a = \theta_Z/\theta_X$.

Lemma 1 stated below provides details about the second order properties of the functions H_Z^{-} and $H_X \circ H_Z^{-}$ (and therefore, on the behavior of the variables Z_i and $\Lambda_X(Z_i)$). These properties not only depend on the position of the parameters τ_X and τ_C with respect to each other, but on their proximity through the parameter r defined by

$$r := 1 - \left|\frac{1}{\tau_C} - \frac{1}{\tau_X}\right| \in [-\infty, 1]$$

(if either $\tau_X = 0$ or $\tau_C = 0$, indeed consider that $r = -\infty$). Its proof can be found in the Supplementary Material document.

Lemma 1. *Let conditions (A_1) and (A_2) hold.*

(i) For different slowly varying functions generically noted v, we have

$$\begin{array}{lll} l_X(x) = c_X(1 - x^{\rho_X}v(x)) & \textit{and} & l_C(x) = c_C(1 - x^{\rho_C}v(x)) \\ \bar{l}_X(x) = c_X^{-\theta_X}(1 - x^{\theta_X\rho_X}v(x)) & \textit{and} & \bar{l}_C(x) = c_C^{-\theta_C}(1 - x^{\theta_C\rho_C}v(x)). \end{array}$$

(ii) The slowly varying functions $\tilde{l}$ and l associated to H_Z^{-} and $H_X \circ H_Z^{-}$ satisfy a second order condition SR2 : as $t \to +\infty$,

$$\frac{\frac{\tilde{l}(tx)}{\tilde{l}(t)} - 1}{\tilde{b}(t)} \longrightarrow K_{\tilde{\rho}}(x) \quad \textit{and} \quad \frac{\frac{l(tx)}{l(t)} - 1}{b(t)} \longrightarrow K_{\rho}(x)$$

where

$$\tilde{\rho} = \rho = \begin{cases} \max(\theta_X\rho_X, -1) & \textit{if } 0 = \tau_X < \tau_C < 1 \\ \max(\theta_C\rho_C, -1) & \textit{if } 0 = \tau_C < \tau_X < 1 \\ 0 & \textit{in the other cases}, \end{cases}$$

and $|\tilde{b}| \in RV_{\tilde{\rho}}$ and $|b| \in RV_{\rho}$. When $\rho = 0$, both $b(t)$ and $\tilde{b}(t)$ are (as $t \to +\infty$) of the order $O((\log t)^{r-1})$ when $r \neq 0$, and of the order $O((\log t)^{-2})$ when $r = 0$.

(iii) The slowly varying function l_Z associated to H_Z satisfies

$$\lim_{x\to+\infty} l_Z(x) = c_Z \begin{cases} \in]0, +\infty[& \textit{if } \tau_X = \tau_C \textit{ or } r \leqslant 0, \\ = +\infty & \textit{if } \tau_X \neq \tau_C \textit{ and } r \in]0,1[\end{cases}$$

where in particular $c_Z = c_X$ if $\tau_X < \tau_C$ and $r < 0$, and $c_Z = c_C$ if $\tau_C < \tau_X$ and $r < 0$. Moreover we have (with the convention $(+\infty)^{-\theta} = 0$ when $\theta > 0$)

$$\tilde{l}(t) \to \tilde{c} := c_Z^{-\theta_Z} \quad \textit{and} \quad l(t) \to c := c_X\tilde{c}^{1/\theta_X}, \quad \textit{as } t \to +\infty.$$

When $\tau_X = \tau_C$ or $r \leqslant 0$, both c and $\tilde{c}$ are positive. When $\tau_X \neq \tau_C$ and $r \in]0,1[$, both $\tilde{c}$ and c are zero and the following relation holds for some $\nu > 0$, as $x \to \infty$

$$\frac{\log l(\exp x)}{x} = -\nu . x^{r-1}(1+o(1)) \longrightarrow 0 \quad \text{and} \quad \frac{\log \tilde{l}(\exp x)}{x} = -\theta_X \nu . x^{r-1}(1+o(1)) \longrightarrow 0 \tag{E.1}$$

Remark 10. *A consequence of this Lemma is that $\bar{l}_X$ and $\bar{l}_C$ also satisfy the SR2 condition with rate functions $|\bar{b}_X| \in RV_{\theta_X \rho_X}$ and $|\bar{b}_C| \in RV_{\theta_C \rho_C}$ respectively.*

Recall now that the function $p(\cdot)$ is defined by

$$p(x) = \mathbb{P}(\delta = 1 | Z = x).$$

The following lemma provides useful developments of functions $p(\cdot)$ and $r(\cdot)$

$$r(t) = p \circ H_Z^{-}(\exp(K_{\tau_Z}(-\log t))),$$

which are crucial to derive the properties of the random proportion $\hat{p}_k$ (and therefore the statements of Proposition 2). Its proof is essantially based on the fact that

$$p(x) = \frac{\bar{F}_C(x) f_X(x)}{\bar{F}_C(x) f_X(x) + \bar{F}_X(x) g(x)} = \left(1 + \frac{(K_{\tau_C}^{-})'(\log H_C(x))}{(K_{\tau_X}^{-})'(\log H_X(x))} \frac{H_C'(x)/H_C(x)}{H_X'(x)/H_X(x)}\right)^{-1}$$

(where f_X and f_C are the respective probability density functions of X and C), as well as on the results of Lemma 1.

Lemma 2. *Let us define the constants*

$$A_X = \theta_X(\tau_X^{-1} - 1)(\tau_X^{-1} + \log c_X) \ , \ A_C = \theta_C(\tau_C^{-1} - 1)(\tau_C^{-1} + \log c_C)$$

and

$$A = A_C - A_X \quad \text{and} \quad B = \frac{\theta_X}{\theta_C} \left(\frac{\tau_X}{\theta_X}\right)^{1-1/\tau_X} \left(\frac{\tau_C}{\theta_C}\right)^{1/\tau_C - 1}.$$

Let assumptions (A_1) and (A_2) hold (the asymptotics below are $x \to +\infty$ and $t \downarrow 0$).

(i) We have

$$p(x) \to p := \begin{cases} 1 & \text{if } 0 \leqslant \tau_X < \tau_C \leqslant 1, \\ 0 & \text{if } 0 \leqslant \tau_C < \tau_X \leqslant 1, \\ \dfrac{\theta_X^{1/\tau_X}}{(\theta_X^{1/\tau_X} + \theta_C^{1/\tau_X})} = a^{1/\tau_X} & \text{if } 0 < \tau_X = \tau_C < 1, \end{cases}$$

and, more precisely,

$$p(x) - p = \begin{cases} D\,(\log x)^{r-1}\left[1 + g(r)(\log x)^{\max(-1,r-1)}(1+o(1))\right] & \text{if } 0 < \tau_X \neq \tau_C \leqslant 1, \\ D\,x^{-1/\theta_X}(\log x)^{\tau_C^{-1}-1}\left[1 + A_C(\log x)^{-1}(1+o(1))\right] & \text{if } 0 = \tau_X < \tau_C \leqslant 1, \\ D\,x^{-1/\theta_C}(\log x)^{\tau_X^{-1}-1}\left[1 + A_X(\log x)^{-1}(1+o(1))\right] & \text{if } 0 = \tau_C < \tau_X \leqslant 1, \\ D\,(\log x)^{-1}(1 + O(1/\log x)) & \text{if } 0 < \tau_C = \tau_X < 1, \end{cases}$$

where

$$D = \begin{cases} -B & \text{if } 0 < \tau_X < \tau_C \leqslant 1, \\ B^{-1} & \text{if } 0 < \tau_C < \tau_X \leqslant 1, \\ -(\tau_C/\theta_C)^{\tau_C^{-1}-1}(\theta_X/\theta_C c_X) & \text{if } 0 = \tau_X < \tau_C \leqslant 1, \\ (\tau_X/\theta_X)^{\tau_X^{-1}-1}(\theta_C/\theta_X c_C) & \text{if } 0 = \tau_C < \tau_X \leqslant 1, \\ -AB(1+B)^{-2} & \text{if } 0 < \tau_C = \tau_X < 1, \end{cases}$$

and

$$g(r) = \begin{cases} A\mathbb{I}_{r<0} + (A-B)\mathbb{I}_{r=0} + (-B)\mathbb{I}_{r\in]0,1[} & \text{if } 0 < \tau_X < \tau_C \leqslant 1, \\ (-A)\mathbb{I}_{r<0} + (-A-B^{-1})\mathbb{I}_{r=0} + (-B^{-1})\mathbb{I}_{r\in]0,1[} & \text{if } 0 < \tau_C < \tau_X \leqslant 1. \end{cases}$$

(ii) When $\tau_Z > 0$ and $\tau_X \neq \tau_C$, as $t \downarrow 0$ we have

$$r(t) - p = D(\theta_Z/\tau_Z)^{r-1}(-\log t)^{-\tau_Z(1-r)}\left(1 + O\left((-\log t)^{-\tau_Z \min\{1,1-r\}}\right)\right),$$

in particular, when $0 < \tau_C < \tau_X \leqslant 1$,

$$r(t) = a^{1/\tau_X}(\tau_X/\tau_C)^{\tau_X^{-1}-1}(-\log t)^{\frac{\tau_C}{\tau_X}-1}\left(1 + O\left((-\log t)^{\max\{-\tau_C,\tau_C/\tau_X-1\}}\right)\right).$$

When $\tau_Z > 0$ *and* $\tau_X = \tau_C$, *we have*

$$r(t) - p = -AB\left[(1+B)^2(\theta_Z/\tau_Z)\right]^{-1}(-\log t)^{-\tau_Z}\left(1 + O\left((-\log t)^{-\tau_Z}\right)\right).$$

When $\tau_Z = 0$, *if* $\tau_+ = \max(\tau_X, \tau_C)$ *we have*

$$r(t) - p = cst(-\log t)^{-1}(\log\log(1/t))^{\frac{1}{\tau_+}-1}\left(1 + O\left((\log\log(1/t))^{-1}\right)\right).$$

with the constant being equal to $\tau_X^{\frac{1}{\tau_X}-1}a^{1/\tau_X}$ *when* $0 = \tau_C < \tau_X \leqslant 1$.

Appendix E.2. Technical Lemmas

The proofs of the following Lemmas can be found in the Supplementary Material document.

Lemma 3. *The denominator of the estimator* $\hat{\theta}_{X,\tau_X}$ *satisfies the relation*

$$D_{k,\tau_X} = \frac{1}{k}\sum_{j=1}^{k} K_{\tau_X}(\hat{\Lambda}_{nX}(Z_{n-j+1,n})) - K_{\tau_X}(\hat{\Lambda}_{nX}(Z_{n-k,n})) = \hat{\Lambda}_k^{\tau_X-1}\hat{p}_k + R_{1,n},$$

where

$$R_{1,n} = \begin{cases} \frac{\tau_X-1}{2}\hat{\Lambda}_k^{\tau_X}\frac{1}{k}\sum_{j=1}^{k}\left(\frac{\hat{\Delta}_{j,k}}{\hat{\Lambda}_k}\right)^2(1+T_{j,k})^{\tau_X-2}, & \text{if } 0 < \tau_X < 1, \\ \frac{1}{k}\sum_{j=1}^{k}\left(\log\left(1+\frac{\hat{\Delta}_{j,k}}{\hat{\Lambda}_k}\right) - \frac{\hat{\Delta}_{j,k}}{\hat{\Lambda}_k}\right) & \text{if } \tau_X = 0, \\ 0 & \text{if } \tau_X = 1 \end{cases}$$

with, for each $j = 1, \ldots, k$, $\hat{\Delta}_{j,k} := \hat{\Lambda}_{nX}(Z_{n-j+1,n}) - \hat{\Lambda}_{nX}(Z_{n-k,n})$ *and the random variable* $T_{j,k}$ *lies between* 0 *and* $\frac{\hat{\Delta}_{j,k}}{\hat{\Lambda}_k}$.

For the following lemma, recall that (E_i) denote the i.i.d. standard exponential variable (E_i) satisfying $Z_i = H_Z^{-}(\exp(K_{\tau_Z}(E_i)))$, and that $l(\cdot)$ denotes the slowly varying function which properties are described in Lemma 1 and which is such that $H_X \circ H_Z^{-}(x) = x^a l(x)$. Note that in part (ii) of this lemma, the results also hold when one replaces $E_{n-k,n}$ by L_{nk}, or replaces $Z_{n-k,n}$ and $E_{n-k,n}$ by $Z_{n-j+1,n}$ and $E_{n-j+1,n}$ (this will occasionally prove useful).

Lemma 4. (i) *For every* $i = 1, \ldots, n$, *and whether* $\tau_Z > 0$ *or is equal to* 0, *we have*

$$\Lambda_X(Z_i) = K_{\tau_X}^{-}\left(aK_{\tau_Z}(E_i) + \log l(\exp K_{\tau_Z}(E_i))\right).$$

(ii) *When* $\tau_Z > 0$, *we have*

$$\Lambda_X(Z_{n-k,n}) = \left(a\frac{\tau_X}{\tau_Z}\right)^{1/\tau_X} E_{n-k,n}^{\tau_Z/\tau_X}(1 + o_{\mathbb{P}}(1)) = \left(a\frac{\tau_X}{\tau_Z}\right)^{1/\tau_X} E_{n-k,n}^{\tau_Z/\tau_X}\left(1 + \beta E_{n-k,n}^{-\alpha}(1 + o_{\mathbb{P}}(1))\right) \quad \text{(E.2)}$$

for some constant β *and exponent* $\alpha = \begin{cases} \tau_Z & \text{when either } \tau_X = \tau_C, \text{ or } \tau_X \neq \tau_C \text{ and } r \leqslant 0, \\ \tau_Z(1-r) & \text{when } \tau_X \neq \tau_C \text{ and } r \in]0,1[. \end{cases}$

When $0 = \tau_X < \tau_C$, *we have* $\Lambda_X(Z_{n-k,n}) = E_{n-k,n}l(E_{n-k,n}) = E_{n-k,n}(1 + o_{\mathbb{P}}(1))$.

When $0 = \tau_C < \tau_X$, *we have*

$$\Lambda_X(Z_{n-k,n}) = (a\tau_X)^{1/\tau_X}(\log E_{n-k,n})^{1/\tau_X}\left(1 + \beta(\log E_{n-k,n})^{-1}(1 + o_{\mathbb{P}}(1))\right).$$

Note that the constant β is negative in the case $\tau_X \neq \tau_C$ and $r \in]0,1[$.

Lemma 5. *Let* $E_1, \ldots, E_n$ *be i.i.d. standard exponential random variables.*

$$\Lambda_k^{\tau_X-1}\hat{p}_k = \begin{cases} \left(\frac{a\tau_X}{\tau_Z}\right)^{1-1/\tau_X} E_{n-k,n}^{\tau_Z(1-1/\tau_X)}\hat{p}_k + R_{2,n}, & \text{if } \tau_X \neq 0 \text{ and } \tau_C \neq 0 \\ \frac{\hat{p}_k}{E_{n-k,n}} + R_{2,n}, & \text{if } 0 = \tau_X < \tau_C < 1 \\ (a\tau_X)^{1-1/\tau_X}(\log(E_{n-k,n}))^{1-1/\tau_X}\hat{p}_k + R_{2,n} & \text{if } 0 = \tau_C < \tau_X < 1, \end{cases}$$

where

$$R_{2,n} = \begin{cases} \left(\frac{a\tau_X}{\tau_Z}\right)^{1-\frac{1}{\tau_X}} E_{n-k,n}^{\tau_Z(1-\frac{1}{\tau_X})} \hat{p}_k \left((1-E_{n-k,n}^{-\tau_Z})^{1-\frac{1}{\tau_X}} \left(1+\frac{1+\tau_X \log l(\exp(K_{\tau_Z}(E_{n-k,n})))}{a\tau_X K_{\tau_Z}(E_{n-k,n})}\right)^{1-\frac{1}{\tau_X}} - 1\right), \\ \quad \text{if } 0<\tau_X<1 \text{ and } \tau_C \neq 0 \\ \frac{\hat{p}_k}{E_{n-k,n}} \left(\frac{1}{l(E_{n-k,n})} - 1\right), \text{ if } 0=\tau_X<\tau_C<1 \\ (a\tau_X)^{1-\frac{1}{\tau_X}} (\log(E_{n-k,n}))^{1-\frac{1}{\tau_X}} \hat{p}_k \left(\left(1+\frac{1+\tau_X \log l(E_{n-k,n})}{a\tau_X \log(E_{n-k,n})}\right)^{1-\frac{1}{\tau_X}} - 1\right), \quad \text{if } 0=\tau_C<\tau_X<1 \\ 0, \quad \text{if } \tau_X = 1 \end{cases}$$

Lemma 6. *Let $E_1, \dots, E_n$ be i.i.d. standard exponential random variables.*

$$\Lambda_k^{-1}\hat{p}_k = \begin{cases} \left(\frac{a\tau_X}{\tau_Z}\right)^{-1/\tau_X} E_{n-k,n}^{-\tau_Z/\tau_X} \hat{p}_k + RR_{2,n}, & \text{if } \tau_X \neq 0 \text{ and } \tau_C \neq 0 \\ \frac{\hat{p}_k}{E_{n-k,n}} + RR_{2,n}, & \text{if } 0=\tau_X<\tau_C<1 \\ (a\tau_X)^{-1/\tau_X} (\log(E_{n-k,n}))^{-1/\tau_X} \hat{p}_k + RR_{2,n} & \text{if } 0=\tau_C<\tau_X<1, \end{cases}$$

where

$$RR_{2,n} = \begin{cases} \left(\frac{a\tau_X}{\tau_Z}\right)^{-\frac{1}{\tau_X}} E_{n-k,n}^{-\frac{\tau_Z}{\tau_X}} \hat{p}_k \left((1-E_{n-k,n}^{-\tau_Z})^{-\frac{1}{\tau_X}} \left(1+\frac{1+\tau_X \log l(\exp(K_{\tau_Z}(E_{n-k,n})))}{a\tau_X K_{\tau_Z}(E_{n-k,n})}\right)^{-\frac{1}{\tau_X}} - 1\right), \\ \quad \text{if } 0<\tau_X<1 \text{ and } \tau_C \neq 0 \\ \frac{\hat{p}_k}{E_{n-k,n}} \left(\frac{1}{l(E_{n-k,n})} - 1\right), \text{ if } 0=\tau_X<\tau_C<1 \\ (a\tau_X)^{-\frac{1}{\tau_X}} (\log(E_{n-k,n}))^{-\frac{1}{\tau_X}} \hat{p}_k \left(\left(1+\frac{1+\tau_X \log l(E_{n-k,n})}{a\tau_X \log(E_{n-k,n})}\right)^{-\frac{1}{\tau_X}} - 1\right), \quad \text{if } 0=\tau_C<\tau_X<1 \\ 0, \quad \text{if } \tau_X = 1 \end{cases}$$

Lemma 7. *Under the assumptions of Theorem 1, we have, as $n \to \infty$,*

$$\begin{array}{ll} \text{if } \tau_X \neq 0 \text{ and } \tau_C \neq 0, & \log(\Lambda_k) = \frac{\tau_Z}{\tau_X} \log L_{nk}(1+o_{\mathbb{P}}(1)) \\ \text{if } \tau_X = 0, & \log(\Lambda_k) = a \log L_{nk}(1+o_{\mathbb{P}}(1)) \\ \text{if } \tau_X \neq 0, \text{ and } \tau_C = 0 & \log(\Lambda_k) = \frac{1}{\tau_X} \log\log L_{nk}(1+o_{\mathbb{P}}(1)) \end{array}$$

Lemma 8. *Under the assumptions of Theorem 4, we have, as n tends to infinity,*

$$\int_{L_k}^{\Lambda_k} u^{\tau_X - 1} \log u \, du = \begin{cases} O_{\mathbb{P}}(\log L_{nk}) & \text{if } \tau_X \neq 0, \tau_C \neq 0 \text{ and } (\tau_X = \tau_C \text{ or } r \leqslant 0), \\ O_{\mathbb{P}}(L_{nk}^{r\tau_Z} \log L_{nk}) & \text{if } \tau_X \neq 0, \tau_C \neq 0, \tau_X \neq \tau_C \text{ and } r \in]0,1[, \\ O_{\mathbb{P}}(\log\log L_{nk}) & \text{if } \tau_X \neq 0 \text{ and } \tau_C = 0, \\ o_{\mathbb{P}}(\log L_{nk}) & \text{if } \tau_X = 0. \end{cases}$$