# A Lynden-Bell integral estimator for extremes of randomly truncated data

Julien Worms, Rym Worms

# A Lynden-Bell integral estimator for extremes of randomly truncated data

J. Worms[1,*], R. Worms[2]

*Keywords:* Extreme values index, Extreme quantiles, Truncated data, Lynden-Bell estimator
*2010 MSC:* 62G32, 62G10

**Abstract**

This work deals with the estimation of the extreme value index and extreme quantiles for heavy tailed data, randomly right truncated by another heavy tailed variable. Under mild assumptions and the condition that the truncated variable is less heavy-tailed than the truncating variable, asymptotic normality is proved for both estimators. The proposed estimator of the extreme value index is an adaptation of the Hill estimator, in the natural form of a Lynden-Bell integral. Simulations illustrate the quality of the estimators under a variety of situations.

## 1. Introduction

Extreme value statistics is an active domain of research, with numerous fields of application, and which benefits from an important litterature in the context of i.i.d. data, dependent data, and (more recently) multivariate or spatial data. For univariate data, semiparametric estimation of the tail of the underlying distribution (for instance, estimation of extreme quantiles) requires in the first place accurate estimation of the so-called extreme-value index (e.v.i.). In the recent years, several authors dedicated their efforts to obtaining good estimations of the e.v.i. for incompletely observed data, *i.e.* randomly censored or truncated data (note here that, since the interest generally lies in the evaluation of the upper tail of the data, left censoring or left truncation is not a relevant framework, and therefore censoring or truncating are considered from the right). In those contexts, the usual estimators of the e.v.i. need some modifications because otherwise they would lead to erroneous estimations when blindly applied to censored or truncated data. Some references for extreme value estimation in the context of randomly censored observations are [1], [2], [3].

The first published work on extreme values estimation under random truncation was written by L.Gardes and G.Stupfler [4], who dealt with heavy-tailed right truncated data (in their work, they provided motivations and many references on main existing results about truncated samples, we refer to [4] in this regard). The framework of randomly right truncated data will be precisely defined in the next section, let us just sketch it for the moment : we consider $\bar{n}$ independent i.i.d. couples $((X_i, Y_i))_{1\leqslant i\leqslant \bar{n}}$ and, among those couples, we only observe those couples which satisfy the condition $X_i \leqslant Y_i$. The actually observed data will then be noted $((X_i^*, Y_i^*))_{1\leqslant i\leqslant n}$. Below, $F$ and $G$ will stand for the respective distributions of $X$ and $Y$, whereas $F^*$ and $G^*$ will stand for the conditional distributions of $X$ and $Y$ given that $X \leqslant Y$ : the latter two are therefore the distributions of the observed samples $(X_i^*)_{1\leqslant i\leqslant n}$ and $(Y_i^*)_{1\leqslant i\leqslant n}$. The first objective is to estimate the e.v.i. of $X$.

The original idea in [4] was to notice that the extreme value indices $\gamma_1^*$ and $\gamma_2^*$ of $F^*$ and $G^*$ are related by a very simple relation to those of $F$ and $G$, $\gamma_1$ and $\gamma_2$ : they proved that we have indeed (when both $F$ and $G$ are heavy-tailed)

$$\gamma_1^* = \gamma_1\gamma_2/(\gamma_1+\gamma_2) \quad \text{and} \quad \gamma_2^* = \gamma_2.$$

*Corresponding author
*Email addresses:* julien.worms@uvsq.fr (J. Worms), rym.worms@u-pec.fr (R. Worms)
[1]Université de Versailles-Saint-Quentin-En-Yvelines, Laboratoire de Mathématiques de Versailles (CNRS UMR 8100), Bât. Fermat, 45 av. des Etats-Unis, 78035 Versailles, France
[2]UPEMLV, UPEC, Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (CNRS UMR 8050), F-94010, Créteil, France

These relations readily yield a proposition of estimator for the parameter of interest $\gamma_1$ by relying on usual Hill estimators of $\gamma_1^*$ and $\gamma_2^*$ :

$$\hat{\gamma}_{GS} = \frac{\hat{\gamma}_1^*(k_1)\hat{\gamma}_2(k_2)}{\hat{\gamma}_2(k_2) - \hat{\gamma}_1^*(k_1)} \quad \text{where} \quad \hat{\gamma}_1^*(k_1) = \frac{1}{k_1}\sum_{i=1}^{k_1} \log \frac{X^*_{n-i+1,n}}{X^*_{n-k_1,n}} \quad \text{and} \quad \hat{\gamma}_2(k_2) = \frac{1}{k_2}\sum_{i=1}^{k_2} \log \frac{Y^*_{n-i+1,n}}{Y^*_{n-k_2,n}} \tag{1}$$

where $X^*_{1,n} \leqslant \ldots \leqslant X^*_{n,n}$ and $Y^*_{1,n} \leqslant \ldots \leqslant Y^*_{n,n}$ denote the usual order statistics of both samples, and $k_1$ and $k_2$ are the number of upper observations which are kept for estimating $\gamma_1^*$ and $\gamma_2^*$.

The authors of [4] also investigated the behavior of an estimator of $F$ in the upper tail, and therefore provided a Weissman-type estimator of extreme quantiles in this truncation context and proved its asymptotic normality. However, their results suffer from some kind of calibration problem, since they are proved only under the condition that one of the numbers $k_1$ and $k_2$ of order statistics used for estimating $\gamma_1^*$ and $\gamma_2$ must grow to infinity faster than the other. The question of getting rid of this restriction was addressed in the prepublication [5].

In this work, we consider the same framework of randomly right-truncated heavy-tailed data, but adopt a new method for defining an estimator of the extreme value index $\gamma_1$ of the truncated sample : in Section 2, this estimator $\widehat{\gamma}_n$ is defined as some Lynden-Bell integral, requiring a single threshold to be chosen, and asymptotic normality is proved for $\widehat{\gamma}_n$ as well as for an estimator of extreme quantiles, under appropriate but mild conditions. Section 3 is devoted to a simulation study illustrating the performance of the defined estimators (with a tentative comparison to the performance of the estimator defined in [4]), and Sections 4 and 5 respectively contain a conclusion and the proofs of the results. The appendix recalls important (and needed) results, previously published in the litterature, and contains as well a technical lemma which is repeatedly used in the proofs section.

## 2. Framework and statement of the results

### *2.1. Notations and definition of the estimators*

Let $((X_i, Y_i))_{1\leqslant i\leqslant \bar{n}}$ be $\bar{n}$ independent copies of a couple $(X, Y)$, where $X$ and $Y$ are positive independent random variables having respective cumulative distribution functions $F$ and $G$. For convenience, we suppose that the lower endpoints of $F$ and $G$ are both equal to 0 (but this will have no influence on the results, since only the highest data values are retained for tail estimation). We assume in this work that $X$ and $Y$ are heavy-tailed distributed, meaning that $1-F$ and $1-G$ (also assumed to be continuous) are regularly varying with respective indices $-1/\gamma_1$ and $-1/\gamma_2$ where $\gamma_1$ and $\gamma_2$ are $> 0$.

We only observe the couples $(X_i, Y_i)$ which satisfy $X_i \leqslant Y_i$ : in other words, the original data $X_i$ are randomly truncated from the right by the $Y_i$, and the actually observed sample is $((X_i^*, Y_i^*))_{1\leqslant i\leqslant N}$, where $N$ follows the $\mathcal{B}(\bar{n}, p)$ distribution, $p$ denoting the (unknown) probability of non-truncation $p = \mathbb{P}(X \leqslant Y)$. Consequently, the distribution of the $X_i^*$ becomes

$$F^*(x) = \mathbb{P}(X \leqslant x | X \leqslant Y) = \frac{1}{p}\int_0^x \overline{G}(t)dF(t). \tag{2}$$

Conditionally on $N = n$, the couples $(X_1^*, Y_1^*), \ldots, (X_N^*, Y_N^*)$ are independent and identically distributed, and $X_i^*$ is no longer independent of $Y_i^*$. It is important to note that, in the sequel, we will work conditionnaly on $N = n$, where $n$ is some deterministic sample size, and we will therefore handle the sample $(X_1^*, Y_1^*), \ldots, (X_n^*, Y_n^*)$ without further reference to $N$.

In this work, $F_n$ will denote the classical Lynden-Bell (nonparametric maximum likelihood) estimator of $F$, namely

$$F_n(x) = \prod_{X_i^* > x} \left(1 - \frac{1}{nC_n(X_i^*)}\right) \quad \text{where} \quad C_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{X_i^* \leqslant x \leqslant Y_i^*}$$

(with the usual convention that a product on the empty set equals 1), where $C_n$ is the estimator of the function $C$

$$C(x) = \mathbb{P}(X \leqslant x \leqslant Y | X \leqslant Y) = p^{-1}\bar{G}(x)F(x) \tag{3}$$

which plays an important role in the analysis of truncated data. Note that $F_n$ is very close to, but different strictly speaking, from the estimator of $F$ considered in [4] ($F_n$ takes rational values, which is not the case of the latter).

Our goal is to adapt the famous Hill estimator in the context of right-truncation. It is well known that (see Remark 1.2.3 in [6] for instance)

$$\mathbb{E}\left[\log(X/t) \,|\, X > t\right] = \frac{1}{\overline{F}(t)} \int_t^\infty \log(x/t)\ dF(x)$$

tends to $\gamma_1$ as $t \to +\infty$. If $(t_n)$ is a sequence of positive thresholds growing to infinity with $n$, we can then define a random version of $\phi(x) = (\overline{F}(t))^{-1} \log(x/t)\mathbb{I}_{x>t}$ by $\hat{\phi}_n(x) = (\overline{F}(t_n))^{-1} \log(x/t_n)\mathbb{I}_{x>t_n}$ and consequently, a natural adaptation of the Hill estimator for $\gamma_1$ is (see relations (1.9) and (1.10) in [7], in the left-truncation case, for details about Lynden-Bell integrals)

$$\widehat{\gamma}_n = \int \hat{\phi}_n(x) dF_n(x) = \frac{1}{n} \sum_{i=1}^n \hat{\phi}_n(X_i^*) \frac{F_n(X_i^*)}{C_n(X_i^*)},$$

which leads to

$$\widehat{\gamma}_n = \frac{1}{n\overline{F}_n(t_n)} \sum_{i=1}^n \log\left(\frac{X_i^*}{t_n}\right) \frac{F_n(X_i^*)}{C_n(X_i^*)} \mathbb{I}_{X_i^* > t_n} \tag{4}$$

Note that this principle has already been successfully applied in the censoring framework in [3] (see equation (7)), where the role of Lynden-Bell estimator was played by the Kaplan-Meier estimator. However, here, the threshold $t_n$ is deterministic instead of being an order statistic. The asymptotic properties of $\widehat{\gamma}_n$ are stated in Theorem 1. Naturally, the lighter the truncation, the closer our estimator $\widehat{\gamma}_n$ gets to the usual Hill estimator. (?)

We will use this estimator of the tail index $\gamma_1$ in order to estimate an extreme quantile, following a classical scheme. More precisely, let $(p_n)$ be some sequence of quantiles orders tending to 0, such that $p_n = o(\overline{F}(t_n))$. If $x_{p_n}$ denote the quantile of $F$ of order $1 - p_n$, *i.e.* solving $\overline{F}(x_{p_n}) = p_n$, then, in this heavy tailed context (see (6) below), it is easy to see that we can define an estimator $\hat{x}_{p_n,t_n}$ of $x_{p_n}$ as

$$\hat{x}_{p_n,t_n} = t_n \left(\frac{\overline{F}_n(t_n)}{p_n}\right)^{\widehat{\gamma}_n}. \tag{5}$$

In the situation of untruncated data, this is a classical estimator for an extreme quantile based on the approximation of the log relative excesses by a Pareto distribution in the heavy-tailed context, where $F_n$ is in this case the empirical distribution function.

### *2.2. Assumptions and results*

The first order condition assumed in this work is the following

$$\overline{F} \in RV_{-1/\gamma_1} \ \text{ and } \ \overline{G} \in RV_{-1/\gamma_2} \ \text{ with } \ 0 < \gamma_1 < \gamma_2. \tag{6}$$

In other words, we assume that the tail of the truncating variable $Y$ is heavier than the tail of the variable $X$ of interest. This condition is needed in many occasions in the proofs of our results, and is due to the presence (in (4)) of the Lynden Bell estimator, evaluated in the tail. Note that this implies the finiteness of the integral $\int_0^\infty dF(x)/\overline{G}(x)$ (which is a sufficient condition sometimes stated in papers dealing with the asymptotic normality of $\overline{F}_n$).

Moreover, if we note $l_F$ the slowly varying function associated to $F$ (*i.e.* such that $\overline{F}(x) = x^{-1/\gamma_1} l_F(x)$), the second order condition we consider is the classical $SR2$ condition for $l_F$ (see [8]),

$$\forall x > 0, \ \frac{l_F(tx)}{l_F(t)} - 1 \ \overset{t\to\infty}{\sim} \ h_{\rho_1}(x)\ g(t) \quad (\forall x > 1) \tag{7}$$

where $g$ is a positive mesurable function, slowly varying with index $\rho_1$, and $h_{\rho_1}(x) = \frac{x^{\rho_1}-1}{\rho_1}$ when $\rho_1 < 0$, or $h_{\rho_1}(x) = \log x$ when $\rho_1 = 0$.

The first assumption on the threshold sequence $(t_n)$ will be that, if we note $\overline{H} = \overline{F}\ \overline{G}$ (note that $H$ is the distribution function of $\min(X, Y)$), $(t_n)$ satisfies

$$n\overline{H}(t_n) \overset{n\to\infty}{\longrightarrow} +\infty. \tag{8}$$

The asymptotic normality result will then require the following condition on $(t_n)$ :

$$\sqrt{n\overline{H}(t_n)}g(t_n) \overset{n\to\infty}{\longrightarrow} \lambda \quad \text{for some } \lambda > 0. \tag{9}$$

**Theorem 1.** *Under assumptions* (6), (7), (8) *and* (9), *as n tends to infinity,*

$$\sqrt{n\overline{H}(t_n)}(\widehat{\gamma}_n - \gamma_1) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\lambda m, s^2\right),$$

*where* $m = \begin{cases} \frac{{\gamma_1}^2}{1-\gamma_1\rho_1} & \text{if } \rho_1 < 0, \\ {\gamma_1}^2 & \text{if } \rho_1 = 0. \end{cases}$ *and* $s^2 = p{\gamma_1}^2\left(1 + \left(\frac{\gamma_1}{\gamma_2}\right)^2\right)\left(1 - \frac{\gamma_1}{\gamma_2}\right)^{-3}$.

Let us now turn to the results about the extreme quantile estimator defined in (5). Suppose that the sequence of quantile orders $(p_n)$, tending to 0, satisfies the condition

$$\overline{F}(t_n)/p_n \overset{n\to\infty}{\longrightarrow} +\infty. \tag{10}$$

**Theorem 2.** *Under (10) and the assumptions of Theorem 1, setting* $d_n = \overline{F}(t_n)/p_n$, *if* $\rho_1 < 0$ *and*

$$\sqrt{n\overline{H}(t_n)}\,/\log d_n \to \infty, \tag{11}$$

*as n tends to* $\infty$ *then*

$$\frac{\sqrt{n\overline{H}(t_n)}}{\log d_n}\left(\frac{\hat{x}_{p_n,t_n}}{x_{p_n}} - 1\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\lambda m, s^2\right)$$

## 3. Finite sample behaviour

In this section, we illustrate our results by presenting some graphics (issued from an extensive study) corresponding to the comparison, in terms of bias and root mean squared error (RMSE), of our new estimator $\widehat{\gamma}_n$ (defined in (4)) with the existing estimator $\hat{\gamma}_{GS}$ (defined in equation (1)) issued from [4], for two classes of heavy-tailed distributions:

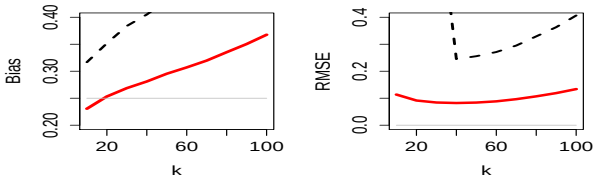
- Burr$(\beta, \tau, \lambda)$ with distribution function $1 - (\frac{\beta}{\beta + x^\tau})^\lambda$, for which the e.v.i. is $\frac{1}{\lambda\tau}$.
- Frechet$(\gamma)$ with distribution function $\exp(-x^{-1/\gamma})$, for which the e.v.i. is $\gamma$.

Note that, in those simulations, we used the random threshold $X^*_{n-k_n,n}$ (where $1 \leqslant k_n < n$) instead of a deterministic threshold $t_n$ in the definition of $\widehat{\gamma}_n$, and we also considered $k_1 = k_2$ in the definition of $\hat{\gamma}_{GS}$, which is out of the scope of Theorem 3 in [4] (but the authors themselves restricted their simulations to this situation, which was then presented as more manageable and convenient). Note that making $n$ vary did not provide notable findings, so we kept the number $n$ of actual observation fixed.

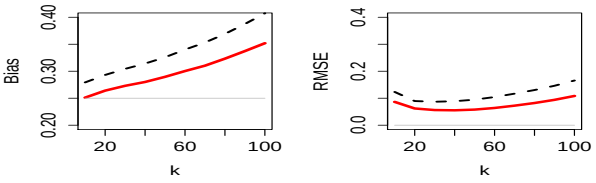
We simulated 2000 random samples of size $n = 200$ in 6 different situations : 3 choices of families of distributions (Burr truncated by another Burr, Fréchet truncated by another Fréchet, and Burr truncated by a Fréchet) combined with 2 choices of truncation strength. This strength is measured by the ultimate probability $\alpha := \frac{\gamma_2}{\gamma_1+\gamma_2}$ of non-truncation in the tail (for a proof of this formula, see [2]), which is distinct from the overall $p = \mathbb{P}(X \leqslant Y)$ : two values were considered, $\alpha = 2/3$ (for $\gamma_1 = 1/4$ and $\gamma_2 = 1/2$, *i.e.* important truncation) and $\alpha = 8/9$ (for $\gamma_1 = 1/4$ and $\gamma_2 = 2$, *i.e.* mild truncation). The results are contained in Figure 1, where bias and RMSE are plotted against different values of $k_n$, the number of excesses used.

This section also contains graphics illustrating the behaviour of our extreme quantile estimator $\hat{x}_{p_n,t_n}$ of $x_{p_n}$ (again computed with the random threshold $X^*_{n-k_n,n}$ instead of $(t_n)$. Under the same simulation framework described above, we considered the estimation of the extreme quantile $x_{p_n}$ with $p_n = 0,03$. Results are displayed in Figure 2.

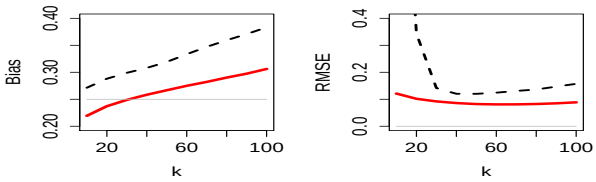
The main conclusion we can deduce from our intensive simulation study is that our estimator $\widehat{\gamma}_n$ seems to behave systematically better (both in terms of bias and RMSE) than the existing estimator $\hat{\gamma}_{GS}$ used with $k_1 = k_2$, whatever the distributions and the value of $\alpha$ are (and changing the sample size yields the same conclusion). Nonetheless, the comparison may be a bit delicate since the properties of $\hat{\gamma}_{GS}$ are only proved when the two numbers $k_1$ and $k_2$



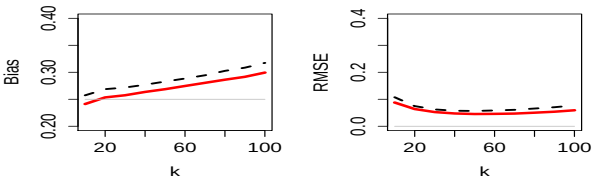

(a) Burr$(10, 4, 1)$ truncated by Burr$(10, 2, 1)$



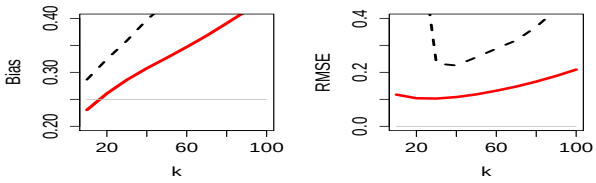

(b) Burr$(10, 4, 1)$ truncated by Burr$(10, 1, 1/2)$



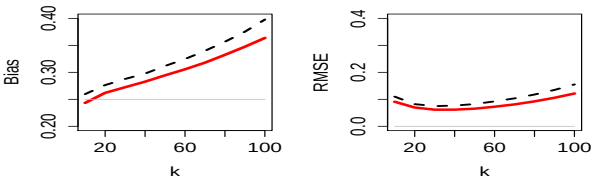

(c) Frechet$(1/4)$ truncated by Frechet$(1/2)$




(d) Frechet$(1/4)$ truncated by Frechet$(2)$




(e) Burr$(10, 4, 1)$ truncated by Frechet$(1/2)$




(f) Burr$(10, 4, 1)$ truncated by Frechet$(2)$

Figure 1: Comparison of bias and RMSE (respectively left and right in each subfigure) for $\hat{\gamma}_n$ (plain) and $\hat{\gamma}_{GS}$ (dashed) where $\gamma_1 = 1/4$, $\gamma_2 = 1/2$ and $\alpha = 2/3$ (important truncation) for subfigures (a),(c),(e), and where $\gamma_1 = 1/4$, $\gamma_2 = 2$ and $\alpha = 8/9$ (mild truncation) for subfigures (b),(d),(f)

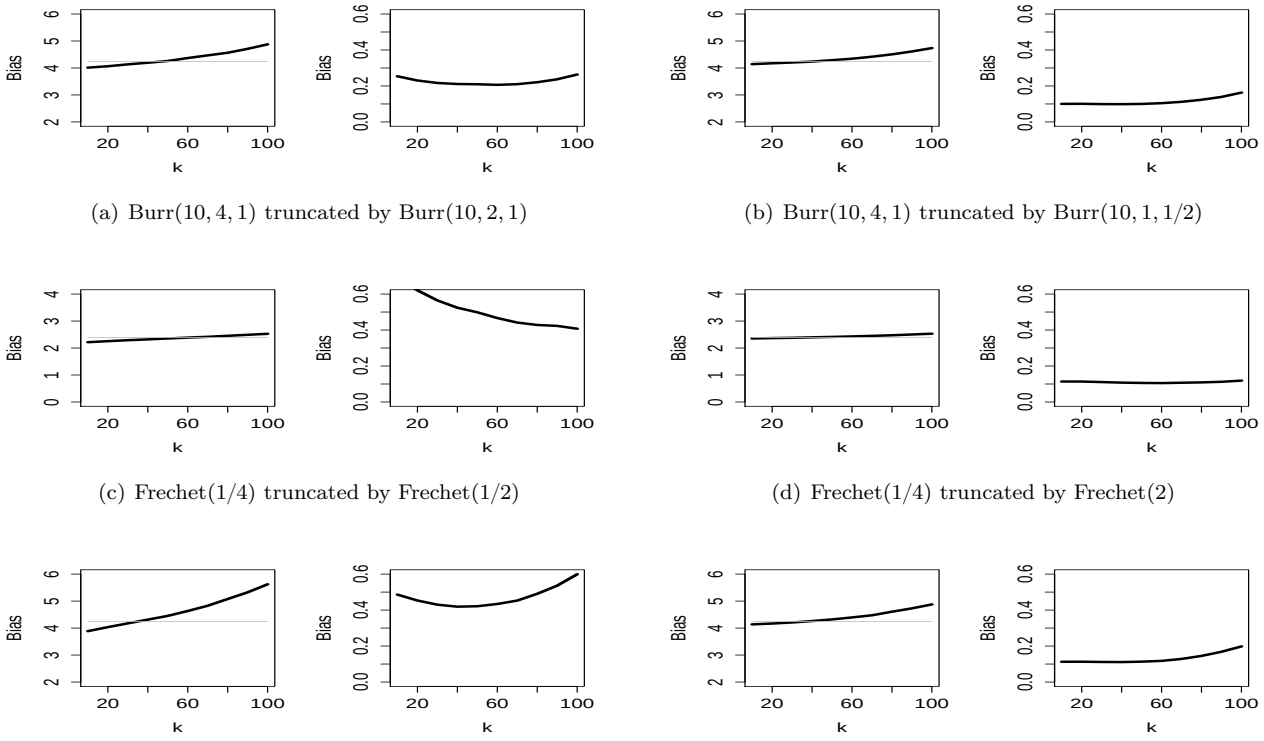
are quite distant from each other. On the other hand, the performance of our estimator clearly diminishes when the ultimate proportion of non-truncation $\alpha$ decreases (which is equivalent to $\gamma_1$ getting closer to $\gamma_2$, which notably increases the asymptotic variance of our estimator) but this phenomenon also holds (and to a greater extent) for $\hat{\gamma}_{GS}$. According to our investigations, and unsurprisingly, a small value of $\rho_1$ also implies a lesser performance. And concerning the bias, since our estimator of $\gamma_1$ is based on the same idea as the Hill estimator in the complete data setting, the relatively high bias observed is neither surprising nor unbearable ; and it is always lower than the bias of $\hat{\gamma}_{GS}$.

Concerning our new extreme quantile estimator $\hat{x}_{p_n,t_n}$, the finite sample behaviour seems quite satisfying, even if its performances depend on the value of $p_n$ and of the truncation strength.

## 4. Conclusion

This paper addressed the problem of estimating tails (extreme value index $\gamma_1$ and extreme quantiles) of randomly right-truncated data, when both the truncated and the truncating variables are heavy-tailed. This framework was first considered in [4], where a first proposition of estimator of $\gamma_1$ was provided. We propose here an alternative approach, leading to an estimator of $\gamma_1$ which takes the form of a Lynden-Bell integral of some particular function, and is therefore a sort of natural version of the Hill estimator in this truncation context. Contrary to the situation of [4] (for which the choice of the numbers of upper order statistics $k_1$ and $k_2$ in the estimator $\hat{\gamma}_{GS}$ defined in (1) could remain very delicate in practice), a single tuning parameter has to be determined (the threshold $t_n$, or in practice the number of upper order statistics), and experimental results are very encouraging.

Concerning the asymptotic normality result for our estimator, the restriction that the truncating variable has a heavier tail than the truncated variable seems to be unavoidable, and improving the performance in term of bias is an open problem, as is the extension of the approach to truncated data with non-negative extreme value index.




(a) Burr$(10, 4, 1)$ truncated by Burr$(10, 2, 1)$

(b) Burr$(10, 4, 1)$ truncated by Burr$(10, 1, 1/2)$

(c) Frechet$(1/4)$ truncated by Frechet$(1/2)$

(d) Frechet$(1/4)$ truncated by Frechet$(2)$

(e) Burr$(10, 4, 1)$ truncated by Frechet$(1/2)$

(f) Burr$(10, 4, 1)$ truncated by Frechet$(2)$

Figure 2: Bias and RMSE (respectively left and right in each subfigure) for $\hat{x}_{p_n,t_n}$ where $\gamma_1 = 1/4$, $\gamma_2 = 1/2$ and $\alpha = 2/3$ (important truncation) for subfigures (a),(c),(e), and where $\gamma_1 = 1/4$, $\gamma_2 = 2$ and $\alpha = 8/9$ (mild truncation) for subfigures (b),(d),(f)

## 5. Proofs of the results

*5.1. Proof of Theorem 1*

We introduce the following important notations : first

$$\widetilde{\gamma}_n = \frac{1}{n}\sum_{i=1}^{n} V_{i,n} \qquad \text{where} \qquad V_{i,n} = \frac{1}{\overline{F}(t_n)} \log\left(\frac{X_i^*}{t_n}\right) \frac{F(X_i^*)}{C(X_i^*)} \mathbb{I}_{X_i^* > t_n} \tag{12}$$

The variables $V_{i,n}$ are independent and identically distributed and, using (2), we readily have $\mathbb{E}(V_{1,n}) = \frac{1}{\overline{F}(t_n)} \int_{t_n}^{\infty} \log(x/t_n) dF(x)$, which converges to $\gamma_1$. Then we consider two (very close but different anyway) estimators of the cumulative hazard function $\Lambda$ of $X$, $\Lambda = -\log F$ : for any $t$, let (for the first definition below, $F_n(t)$ is supposed $> 0$ though)

$$\Lambda_n(t) = -\log F_n(t) \qquad \text{and} \qquad \hat{\Lambda}_n(t) = \sum_{X_i^* > t} \frac{1}{nC_n(X_i^*)}. \tag{13}$$

We will later approach $\hat{\Lambda}_n(t_n)/\overline{F}(t_n)$ by $\frac{1}{n}\sum_{i=1}^{n} V'_{i,n}$, where the i.i.d. variables $V'_{i,n}$ are defined by

$$V'_{i,n} = \frac{\mathbb{I}_{X_i^* > t_n}}{\overline{F}(t_n) C(X_i^*)} \qquad \text{with} \qquad \mathbb{E}(V'_{1,n}) = \frac{\Lambda(t_n)}{\overline{F}(t_n)}. \tag{14}$$

Finally we set $W_{i,n} = V_{i,n} - \mathbb{E}(V_{1,n})$ and $W'_{i,n} = V'_{i,n} - \mathbb{E}(V'_{1,n})$, as well as

$$\Delta_n = \overline{F}_n(t_n)/\overline{F}(t_n) \quad \text{and} \quad v_n = n\overline{H}(t_n)$$

Before proceeding to the proof of Theorem 1, let us state some lemmas (complèter bien sûr les conditions/hypothèses...)

**Lemma 1.** *Under condition* (6)*, we have* $\Delta_n \widehat{\gamma}_n - \widetilde{\gamma}_n = o_{\mathbb{P}}(v_n^{-1/2})$.

**Lemma 2.** *Under conditions* (8) *and* (6)*, the sequence* $(\Delta_n)$ *converges to* 1 *in probability.*

**Lemma 3.** *If* $T = \max\{X_i^*; nC_n(X_i^*) = 1\}$ *and* $A_n = \{T \leqslant t_n\}$*, then, under condition* (6)*, we have*

$$\frac{\sqrt{v_n}}{\overline{F}(t_n)} \mathbb{I}_{A_n} (\Lambda_n(t_n) - \hat{\Lambda}_n(t_n)) = o_{\mathbb{P}}(1).$$

**Lemma 4.** *Under conditions* (8) *and* (6)*,*

$$\sqrt{v_n} \frac{\hat{\Lambda}_n(t_n) - \Lambda(t_n)}{\overline{F}(t_n)} = \sqrt{v_n} \overline{W}'_n + o_{\mathbb{P}}(1). \tag{15}$$

For the next two lemmas, note that quantities $s^2$ and $m$ have been defined in the statement of Theorem 1).

**Lemma 5.** *Under conditions* (8) *and* (6)*, the sequences* $\sqrt{v_n}\overline{W}_n$, $\sqrt{v_n}\overline{W}'_n$ *and* $\sqrt{v_n}(\overline{W}_n - \gamma_1 \overline{W}'_n)$ *converge in distribution to centered gaussian distributions of respective variances* $2p{\gamma_1}^2/(1-\gamma_1/\gamma_2)^3$, $p/(1-\gamma_1/\gamma_2)$ *and* $s^2$.

**Lemma 6.** *Under conditions* (7) *and* (9)*, we have* $\sqrt{v_n}(\mathbb{E}(\widetilde{\gamma}_n) - \gamma_1) \overset{n\to\infty}{\longrightarrow} \lambda m$.

Note that Lemma 2 is a direct corollary of relation (17) and of Lemmas 4 and 5. Lemma 4 is included in the proof of Theorem 1 in [4]. We will provide the proofs of the other lemmas in the next subsections.

Let us now turn to the proof of Theorem 1. We have, thanks to Lemmas 1 and 2,

$$\sqrt{v_n}(\widehat{\gamma}_n - \gamma_1) \;=\; \sqrt{v_n}(\Delta_n^{-1}\widetilde{\gamma}_n - \gamma_1) + o_{\mathbb{P}}(1) \;=\; \Delta_n^{-1}\sqrt{v_n}\,(\,(\widetilde{\gamma}_n - \gamma_1) - \gamma_1(\Delta_n - 1)\,) + o_{\mathbb{P}}(1). \tag{16}$$

We consider

$$\Delta_n - 1 = \frac{\overline{F}_n(t_n) - \overline{F}(t_n)}{\overline{F}(t_n)} = -\frac{F_n(t_n) - F(t_n)}{\overline{F}(t_n)}$$

and we want to deal with this difference by introducing cumulative hazard functions (defined at the beginning of this section). But if there exists some data value $X_i^*$ which is both greater than $t_n$ and such that $nC_n(X_i^*) = 1$, then $F_n(t_n) = 0$ and $\Lambda_n(t_n)$ is undefined. In order to avoid this, we introduce the variable

$$T = \max\{X_i^*; nC_n(X_i^*) = 1\}$$

for which [9] proved that $\mathbb{P}(T = \min_{i\leqslant n} X_i^*)$ converges to 1. Therefore, if we set $A_n = \{T \leqslant t_n\}$, then on $A_n$ we have $F_n(t_n) > 0$ on one hand, and on the other hand $\mathbb{P}(A_n^c) \leqslant \mathbb{P}(T \neq \min_{i\leqslant n} X_i^*) + \mathbb{P}(\min_{i\leqslant n} X_i^* > t_n)$, which tends to 0. We can thus write, using the mean value theorem,

$$\begin{aligned}\Delta_n - 1 &= -\frac{\exp(-\Lambda_n(t_n)) - \exp(-\Lambda(t_n))}{\overline{F}(t_n)}\mathbb{I}_{A_n} + \frac{\overline{F}_n(t_n) - \overline{F}(t_n)}{\overline{F}(t_n)}\mathbb{I}_{A_n^c} \\ &= \xi_n\, \mathbb{I}_{A_n}\frac{\Lambda_n(t_n) - \Lambda(t_n)}{\overline{F}(t_n)} + \frac{\overline{F}_n(t_n) - \overline{F}(t_n)}{\overline{F}(t_n)}\mathbb{I}_{A_n^c}\end{aligned}$$

where $\xi_n$ converges to 1 in probability, since both $\Lambda_n(t_n)$ and $\Lambda(t_n)$ converge to 0. Therefore, using successively $\mathbb{P}(A_n^c) \to 0$ and Lemmas 3, 4 and 5, we can write

$$\begin{aligned}\sqrt{v_n}(\Delta_n - 1) &= \xi_n \mathbb{I}_{A_n}\sqrt{v_n}\frac{\hat{\Lambda}_n(t_n) - \Lambda(t_n)}{\overline{F}(t_n)} + o_{\mathbb{P}}(1) \;=\; \xi_n \mathbb{I}_{A_n}\sqrt{v_n}\overline{W}'_n + o_{\mathbb{P}}(1) \\ &= \sqrt{v_n}\overline{W}'_n + o_{\mathbb{P}}(1). \qquad (17)\end{aligned}$$

On the other hand,

$$\sqrt{v_n}(\widetilde{\gamma}_n - \gamma_1) = \sqrt{v_n}\overline{W}_n + \sqrt{v_n}(\mathbb{E}(\widetilde{\gamma}_n) - \gamma_1)$$

and consequently, combining relations (16) and (17) with Lemmas 5 and 6, the theorem is proved :

$$\sqrt{v_n}(\widehat{\gamma}_n - \gamma_1) \;=\; \Delta_n^{-1}\left\{\sqrt{v_n}(\overline{W}_n - \gamma_1\overline{W}'_n) + \sqrt{v_n}(\mathbb{E}(\widetilde{\gamma}_n) - \gamma_1) + o_{\mathbb{P}}(1)\right\} + o_{\mathbb{P}}(1) \;\xrightarrow{\mathcal{L}}\; \mathcal{N}\left(\lambda m, s^2\right).$$

*5.2. Proof of Theorem 2*

Recall that $d_n = \frac{\overline{F}(t_n)}{p_n} \to \infty$, and the notations $\Delta_n = \frac{\overline{F}_n(t_n)}{\overline{F}(t_n)}$ (which satisfies (17)) and $v_n = n\overline{H}(t_n)$. We write

$$\frac{\hat{x}_{p_n,t_n}}{x_{p_n}} - 1 \;=\; \frac{t_n}{x_{p_n}}(\Delta_n d_n)^{\widehat{\gamma}_n} - 1 \;=\; \Delta_n^{\widehat{\gamma}_n}\left(\frac{t_n}{x_{p_n}}\, d_n^{\gamma_1}T_n^1 + T_n^2 + T_n^3\right),$$

where $T_n^1 := d_n^{\widehat{\gamma}_n - \gamma_1} - 1$, $T_n^2 := \frac{t_n}{x_{p_n}}d_n^{\gamma_1} - 1$ and $T_n^3 := 1 - \Delta_n^{-\widehat{\gamma}_n}$. We are going to prove that both $T_n^2$ and $T_n^3$ are $o_{\mathbb{P}}(\log d_n/\sqrt{v_n})$, and that $\frac{\sqrt{v_n}}{\log d_n}T_n^1 \xrightarrow{\mathcal{L}} \mathcal{N}\left(\lambda m, s^2\right)$. This will conclude the proof, since both $\Delta_n$ and $\frac{t_n}{x_{p_n}}\, d_n^{\gamma_1}$ tend to 1.

Let us first focus on $T_n^1$. The mean value theorem yields

$$\frac{\sqrt{v_n}}{\log d_n}T_n^1 = \sqrt{v_n}(\widehat{\gamma}_n - \gamma_1)\exp(E_n),$$

where $|E_n| \leqslant |\widehat{\gamma}_n - \gamma_1|\log d_n$ and therefore $E_n$ tends to 0 thanks to Theorem 1 and assumption (11). The desired result for $T_n^1$ is then implied by Theorem 1.

We now deal with $T_n^2$. Recalling that $\overline{F}(x) = x^{-\gamma_1}l_F(x)$, by definition of $x_{p_n}$ we have

$$T_n^2 = \left(\frac{l_F(x_{p_n})}{l_F(t_n)}\right)^{-\gamma_1} - 1$$

We use the following representation of $l_F$ (see [10] page 1195) when $\rho < 0$ :

$$l_F(x) = C\left(1 + {\rho_1}^{-1}g(x) + o(g(x))\right), \text{ for } x \to +\infty.$$

Hence

$$\frac{l_F(x_{p_n})}{l_F(t_n)} = 1 - {\rho_1}^{-1}g(t_n)\left(1 - \frac{g(x_{p_n})}{g(t_n)} + o_{\mathbb{P}}(1) + o\left(\frac{g(x_{p_n})}{g(t_n)}\right)\right).$$

But $g(x_{p_n})/g(t_n)$ tends to 0 because $x_{p_n}/t_n$ tends to infinity and

$$|\, g(x_{p_n})/g(t_n) \,-\, (x_{p_n}/t_n)^{\rho_1}\,| \;\leqslant\; \sup_{y\geqslant 1}\left|g(yt_n)/g(t_n) - y^{-\rho}\right| \;\longrightarrow\; 0.$$

It follows that $\frac{l_F(x_{p_n})}{l_F(t_n)} = 1 - {\rho_1}^{-1} g(t_n)(1+o_{\mathbb{P}}(1))$. Thus $\left|(l_F(x_{p_n})/l_F(t_n))^{-\gamma_1} - 1\right| \leqslant c\,|l_F(x_{p_n})/l_F(t_n) - 1|$, for some constant $c$ and then

$$\frac{\sqrt{v_n}}{\log d_n}|T_n^2| \leqslant c{\rho_1}^{-1}\sqrt{v_n}g(t_n)\frac{1+o_{\mathbb{P}}(1)}{\log d_n}.$$

Assumption (9) and the fact that $\log d_n$ tends to 0 conclude the proof for $T_n^2$.

For $T_n^3$, we use the mean value theorem to write

$$T_n^3 = \widehat{\gamma}_n K_n^{-\widehat{\gamma}_n - 1}(\Delta_n - 1),$$

with $K_n$ tending to 1. In view of (17) and Lemma 5 , we thus have $\frac{\sqrt{v_n}}{\log d_n}(\Delta_n - 1) = O_{\mathbb{P}}(1)/\log d_n = o_{\mathbb{P}}(1)$ and then the desired neglibility of $T_n^3$ follows.

*5.3. Proof of Lemma 1*

We have $\Delta_n\widehat{\gamma}_n = \widetilde{\gamma}_n + S_{n,1} + S_{n,2}$, with

$$S_{n,1} := \frac{1}{\overline{F}(t_n)}\frac{1}{n}\sum_{i=1}^{n}\frac{F_n(X_i^*) - F(X_i^*)}{C_n(X_i^*)}\,\log\left(\frac{X_i^*}{t_n}\right)\mathbb{I}_{X_i^*>t_n}$$

and

$$S_{n,2} := \frac{1}{\overline{F}(t_n)}\frac{1}{n}\sum_{i=1}^{n}F(X_i^*)\left(\frac{1}{C_n(X_i^*)} - \frac{1}{C(X_i^*)}\right)\,\log\left(\frac{X_i^*}{t_n}\right)\mathbb{I}_{X_i^*>t_n}.$$

Let us show that both $\sqrt{v_n}S_{n,1}$ and $\sqrt{v_n}S_{n,2}$ are $o_{\mathbb{P}}(1)$. On one hand,

$$|\sqrt{v_n}S_{n,1}| \leqslant \left(\sqrt{n}\sup_{x>t_n}|F_n(x) - F(x)|\right)\,\sup_{X_i^*>t_n}\frac{C(X_i^*)}{C_n(X_i^*)}\,\sqrt{\overline{H}(t_n)}\bar{V}_n^1 \tag{18}$$

where $\bar{V}_n^1 := \frac{1}{n}\sum_{i=1}^{n}V_{i,n}^1$ with

$$V_{i,n}^1 := \frac{1}{\overline{F}(t_n)}\frac{\mathbb{I}_{X_i^*>t_n}}{C(X_i^*)}\,\log\left(\frac{X_i^*}{t_n}\right).$$

Using (2) and (3) yields

$$\mathbb{E}(V_{i,n}^1) = \frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\frac{1}{F(x)}\log(x/t_n)dF(x) = (1+o_{\mathbb{P}}(1))\frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\log(x/t_n)dF(x),$$

which converges to $\gamma_1$ ; Markov inequality then yields $\sqrt{\overline{H}(t_n)}\bar{V}_n^1 = o_{\mathbb{P}}(1)$. On the other hand,

$$|\sqrt{v_n}S_{n,2}| \leqslant \sup_{X_i^*>t_n}\frac{C(X_i^*)}{C_n(X_i^*)}\,\left(\sqrt{n}\sup_{X_i^*>t_n}|C_n(X_i^*) - C(X_i^*)|\right)\,\sqrt{\overline{H}(t_n)}\bar{V}_n^2, \tag{19}$$

where $\bar{V}_n^2 := \frac{1}{n}\sum_{i=1}^{n}V_{i,n}^2$ with

$$V_{i,n}^2 := \frac{1}{\overline{F}(t_n)}\frac{F(X_i^*)}{C^2(X_i^*)}\,\log\left(\frac{X_i^*}{t_n}\right)\mathbb{I}_{X_i^*>t_n}.$$

Using again (2) and (3), we have

$$\mathbb{E}(V_{i,n}^2) = p\frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\frac{\log(x/t_n)}{F(x)\overline{G}(x)}dF(x) = p(1+o_{\mathbb{P}}(1))\frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\frac{1}{\overline{G}(x)}\log(x/t_n)dF(x).$$

By Lemma 8 (where constant $c_1$ is defined), it comes $\mathbb{E}(V_{i,n}^2) = (1+o_{\mathbb{P}}(1))\frac{pc_1}{\overline{G}(t_n)}$ and Markov inequality then yields $\sqrt{\overline{H}(t_n)}\bar{V}_n^2 = O_{\mathbb{P}}\left((\overline{F}(t_n)/\overline{G}(t_n))^{1/2}\right) = o_{\mathbb{P}}(1)$. Combining (18) and (19) with Lemma 7 ends the proof.

*5.4. Proof of Lemma 3*

Recall that $T = \max\{X_i^*; nC_n(X_i^*) = 1\}$ and that we previously saw that $\mathbb{P}(A_n) \to 1$ when $A_n = \{T \leqslant t_n\}$. Using the fact that $0 \leqslant -\log(1-x) - x \leqslant \frac{x^2}{1-x}$ for any $0 \leqslant x < 1$, and that, on $A_n$, we have $nC_n(X_i^*) \geqslant 2$ for every $X_i^* > t_n$, we can write that

$$\frac{\sqrt{v_n}}{\overline{F}(t_n)}\mathbb{I}_{A_n}\ |\Lambda_n(t_n) - \hat{\Lambda}_n(t_n)| \ \leqslant\ \frac{\sqrt{v_n}}{\overline{F}(t_n)}\mathbb{I}_{A_n}\sum_{X_i^*>t_n}\frac{1}{n^2C_n^2(X_i^*)}\ \frac{1}{1-\frac{1}{nC_n(X_i^*)}}\ \leqslant\ 2\mathbb{I}_{A_n}\frac{\sqrt{v_n}}{\overline{F}(t_n)}\ \sum_{X_i^*>t_n}\frac{1}{n^2C_n^2(X_i^*)}$$

Using Lemma 7, we have

$$\frac{\sqrt{v_n}}{\overline{F}(t_n)}\ \sum_{X_i^*>t_n}\frac{1}{n^2C_n^2(X_i^*)} \leqslant O_{\mathbb{P}}(1)\sqrt{\frac{\overline{G}(t_n)}{n\overline{F}(t_n)}}\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{I}_{X_i^*>t_n}}{C^2(X_i^*)}.$$

Noting $Z_n = \frac{1}{n}\sum_{i=1}^n \mathbb{I}_{X_i^*>t_n}/C^2(X_i^*)$, and using (2) and (3), we have

$$\mathbb{E}(Z_n) = \int_{t_n}^{\infty}\frac{p}{F^2(x)}\frac{dF(x)}{\overline{G}(x)} = p(1+o_{\mathbb{P}}(1))\int_{t_n}^{\infty}\frac{dF(x)}{\overline{G}(x)}.$$

Via Lemma 8, $\mathbb{E}\left(\sqrt{\frac{\overline{G}(t_n)}{n\overline{F}(t_n)}}Z_n\right)$ tends to 0 and therefore $\sqrt{\frac{\overline{G}(t_n)}{n\overline{F}(t_n)}}Z_n = o_{\mathbb{P}}(1)$ by Markov's inequality, which ends the proof of the lemma.

*5.5. Proof of Lemma 5*

For brevity, we only prove the third part of the lemma. First, using relation (2) and Lemma 8 (wherein the constants $c_0 = 1/q$, $c_1 = \gamma_1/q^2$, $c_2 = 2{\gamma_1}^2/q^3$ are defined, with $q = 1 - \gamma_1/\gamma_2$), it is easily seen that

$$\begin{aligned}
\mathbb{E}(V_{1,n}) &= \frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\log\left(\frac{x}{t_n}\right)dF(x) \overset{n\to\infty}{\longrightarrow} \gamma_1\\
\mathbb{E}(V_{1,n}^2) &= \frac{p}{{\overline{F}(t_n)}^2}\int_{t_n}^{\infty}\log^2\left(\frac{x}{t_n}\right)\frac{dF(x)}{\overline{G}(x)} = \frac{pc_2}{\overline{H}(t_n)}(1+o(1))\\
\mathbb{E}(V'_{1,n}) &= \frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\frac{dF(x)}{F(x)} = \frac{\Lambda(t_n)}{\overline{F}(t_n)} \overset{n\to\infty}{\longrightarrow} 1\\
\mathbb{E}((V'_{1,n})^2) &= \frac{1}{{\overline{F}(t_n)}^2}\int_{t_n}^{\infty}\frac{p}{\overline{G}(x)F^2(x)}dF(x) = \frac{p(1+o(1))}{{\overline{F}(t_n)}^2}\int_{t_n}^{\infty}\frac{dF(x)}{\overline{G}(x)} = \frac{pc_0}{\overline{H}(t_n)}(1+o(1))\\
\mathbb{E}(V_{1,n}V'_{1,n}) &= \frac{p(1+o(1))}{{\overline{F}(t_n)}^2}\int_{t_n}^{\infty}\log\left(\frac{x}{t_n}\right)\frac{dF(x)}{\overline{G}(x)} = \frac{pc_1}{\overline{H}(t_n)}(1+o(1))
\end{aligned}$$

Introducing $U_{i,n} = W_{i,n} - \gamma_1 W'_{i,n}$ and $S_n = \sum_{i\leqslant n} U_{i,n}$, we thus obtain ($s^2$ is defined in the statement of the lemma)

$$\mathbb{V}ar(U_{1,n}) = \mathbb{V}ar(V_{1,n} - \gamma_1 V'_{1,n}) = \frac{s^2}{\overline{H}(t_n)}(1+o(1))$$

and consequently $\sqrt{v_n}(\overline{W}_n - \gamma_1\overline{W}'_n) = \sqrt{v_n}S_n/n = s(1+o(1))S_n/\mathbb{V}ar(S_n)$, which converges in distribution to $\mathcal{N}(0,s^2)$ as soon as Lyapunov's condition holds. After some simplifications, Lyapunov's condition becomes the existence of some $\delta > 0$ such that

$$n^{-\delta/2}(\overline{H}(t_n))^{1+\delta/2}\mathbb{E}(|U_{1,n}|^{2+\delta}) \overset{n\to\infty}{\longrightarrow} 0.$$

Proceeding as in [4], and noting that $\mathbb{E}(V_{1,n}) - \gamma_1\mathbb{E}(V'_{1,n})$ vanishes to 0, the double application of the inequality $|a+b|^{2+\delta} \leqslant 2^{1+\delta}(|a|^{2+\delta} + |b|^{2+\delta})$ shows that it suffices to prove the following, for some $\delta > 0$ :

$$n^{-\delta/2}(\overline{H}(t_n))^{1+\delta/2}\mathbb{E}(|V|^{2+\delta}) \overset{n\to\infty}{\longrightarrow} 0 \quad \text{for both } V = V_{1,n} \text{ and } V = V'_{1,n} \tag{20}$$

We prove this property for $V = V_{1,n}$, the proof for $V = V'_{1,n}$ being very similar. We have

$$\mathbb{E}(|V_{1,n}|^{2+\delta}) \;=\; p^{2+\delta}(\overline{F}(t_n))^{-2-\delta}\int_{t_n}^{\infty}\log^{2+\delta}\left(\frac{x}{t_n}\right)\frac{dF(x)}{\overline{G}^{1+\delta}(x)}$$

Mimicking the proof of Lemma 8 stated in the appendix, and because $\delta$ can be chosen arbitrary small (so that $(1+\delta)/\gamma_2$ remains lower than $1/\gamma_1$), we can prove that

$$\frac{\overline{G}^{1+\delta}(t_n)}{\overline{F}(t_n)}\int_{t_n}^{\infty}\log^{2+\delta}\left(\frac{x}{t_n}\right)\frac{dF(x)}{\overline{G}^{1+\delta}(x)} \;=\; O(1)$$

and therefore, since we assumed that $n\overline{H}(t_n)\to\infty$, the desired property (20) holds for $V = V_{1,n}$ :

$$n^{-\delta/2}(\overline{H}(t_n))^{1+\delta/2}\mathbb{E}(|V_{1,n}|^{2+\delta}) \;\leqslant\; O(1)n^{-\delta/2}(\overline{H}(t_n))^{1+\delta/2}(\overline{F}(t_n))^{-2-\delta}\overline{F}(t_n)\overline{G}^{-1-\delta}(t_n) \;=\; O(1)(n\overline{H}(t_n))^{-\delta/2} \overset{n\to\infty}{\longrightarrow} 0.$$

*5.6. Proof of Lemma 6*

Recall that $\mathbb{E}(\widetilde{\gamma}_n) = \frac{1}{\overline{F}(t_n)}\int_{t_n}^{\infty}\log\left(\frac{x}{t_n}\right)dF(x) = \int_1^{+\infty}\frac{1}{y}\,\frac{\overline{F}(yt_n)}{\overline{F}(t_n)}dy$ by integration by parts and change of variables. Since $\overline{F}(y) = y^{-1/\gamma_1}l_F(y)$, we have

$$\sqrt{v_n}(\mathbb{E}(\widetilde{\gamma}_n) - \gamma_1) = \sqrt{v_n}\int_1^{+\infty} y^{-1/\gamma_1-1}\left(\frac{l_F(yt_n)}{l_F(t_n)} - 1\right)\,dy,$$

and using assumption (7) and Proposition 3.1 in [10], we can write

$$\int_1^{+\infty} y^{-1/\gamma_1-1}\left(\frac{l_F(yt_n)}{l_F(t_n)} - 1\right)\,dy = g(t_n)\;\int_1^{+\infty} y^{-1/\gamma_1-1}h_{\rho_1}(y)dy + o(g(t_n)).$$

The result then follows from assumption (9) and the fact that $\int_1^{+\infty} y^{-1/\gamma_1-1}h_{\rho_1}(y)dy \;=\; m$.

## 6. Appendix

This appendix contains two lemmas : Lemma 7 contains results which are proved elsewhere but are crucial for our proof, and which we thus restate here, whereas Lemma 8 is a variant of a particular case of Lemma 2 in [4], and states essential equivalences for our proofs.

**Lemma 7.** *If $t_n$ tends to infinity with $n$, then*

*(a)* $\sqrt{n}\sup_{x>t_n}|F_n(x) - F(x)| = O_{\mathbb{P}}(1)$.

*(b)* $\sup_{1\leqslant i\leqslant n}\left\{\frac{C(X_i^*)}{C_n(X_i^*)}\,\middle|\, X_i^* > t_n\right\} = O_{\mathbb{P}}(1)$.

*(c)* $\sqrt{n}\sup_{1\leqslant i\leqslant n}\left\{\,|C_n(X_i^*) - C(X_i^*)|\;\middle|\; X_i^* > t_n\,\right\} = O_{\mathbb{P}}(1)$.

*Proof*

$(a)$ is a consequence of point 6 page 176 in [11]. $(b)$ is proved in [4] (see lemma 5), following the ideas contained in [12]. Since $C_n = F_n^* - G_n^*$ , where $F_n^*$ and $G_n^*$ are respectively the empirical distribution functions of $F^*$ and $G^*$, $(c)$ is a consequence of $\sqrt{n}\sup_{x\geqslant 0}|F_n^*(x) - F^*(x)| = O_{\mathbb{P}}(1)$ and $\sqrt{n}\sup_{x\geqslant 0}|G_n^*(x) - G^*(x)| = O_{\mathbb{P}}(1)$ (see [11] pages 172-173).

**Lemma 8.** *Under condition* (6)*, for any $k\in\mathbb{N}$, as $n\to\infty$,*

$$\int_{t_n}^{\infty}\log^k\left(\frac{x}{t_n}\right)\,\frac{dF(x)}{\overline{G}(x)} = c_k\frac{\overline{F}(t_n)}{\overline{G}(t_n)}(1+o(1))$$

*where* $c_k = \dfrac{{\gamma_1}^k\,k!}{(1-\gamma_1/\gamma_2)^{k+1}}$.

*Proof*

Let us note $\alpha = 1/\gamma_2$ and $\beta = 1/\gamma_1$, which satisfy $0 < \alpha < \beta$ by assumption. We need to prove that the following quantity converges to $c_k$ (below, $\delta > 0$ is arbitrary small)

$$\begin{aligned}
&\frac{\overline{G}(t_n)}{\overline{F}(t_n)}\int_{t_n}^{\infty}\log^k\left(\frac{x}{t_n}\right)\frac{dF(x)}{\overline{G}(x)} \\
&\quad = -\int_1^{\infty}\log^k(y)\frac{\overline{G}(t_n)}{\overline{G}(yt_n)}\frac{t_n d\overline{F}(yt_n)}{\overline{F}(t_n)} \\
&\quad = -\int_1^{\infty}\log^k(y)\; y^{\alpha}\;\frac{t_n d\overline{F}(yt_n)}{\overline{F}(t_n)} \\
&\qquad\qquad -\int_1^{\infty}\log^k(y) y^{\alpha+\delta}\left\{\frac{\overline{G}(t_n)}{\overline{G}(yt_n)}\frac{(yt_n)^{-\alpha-\delta}}{t_n^{-\alpha-\delta}} - y^{-\delta}\right\}\frac{t_n d\overline{F}(yt_n)}{\overline{F}(t_n)} \\
&\quad = I_{n,k}(\alpha) + o(1)I_{n,k}(\alpha+\delta) \qquad (21)
\end{aligned}$$

In the last line, we used Theorem 1.5.2 in [8] with the fact that $x \mapsto x^{-\alpha-\delta}/\overline{G}(x)$ is regularly varying of order $-\delta$. It thus remains to prove that $I_{n,k}(\alpha)$ converges to $c_k$ (the same being true for $I_{n,k}(\alpha+\delta)$). We now introduce the notations : for $\theta > 0$

$$J_k(\theta) = \int_1^{\infty}\log^k(y)y^{-\theta-1}dy = \frac{k!}{\theta^{k+1}} \qquad \text{and} \qquad J_{n,k} = \int_1^{\infty}\log^k(y)y^{\alpha-1}\frac{\overline{F}(yt_n)}{\overline{F}(t_n)}dy.$$

For any $\delta \in ]0, \beta-\alpha[$, since the function $x \mapsto x^{\beta-\delta}\overline{F}(x)$ is regularly varying of order $-\delta$, we have

$$\begin{aligned}
J_{n,k} &= \int_1^{\infty}\log^k(y)y^{\alpha-\beta-1}dy \;+\; \int_1^{\infty}\log^k(y)y^{\alpha-1}\left(\frac{\overline{F}(yt_n)}{\overline{F}(t_n)}\frac{(yt_n)^{\beta-\delta}}{t_n^{\beta-\delta}} - y^{-\delta}\right)y^{-\beta+\delta}dy \\
&= J_k(\beta-\alpha) + o(1)
\end{aligned}$$

We thus have, by integration by parts and the relation $kJ_{k-1}(\theta) = \theta J_k(\theta)$,

$$\begin{aligned}
I_{n,k}(\alpha) &= \int_1^{\infty}(k\log^{k-1}(y) + \alpha\log^k(y))y^{\alpha-1}\frac{\overline{F}(yt_n)}{\overline{F}(t_n)}dy \\
&= k\,J_{n,k-1} + \alpha\,J_{n,k} \\
&\overset{n\to\infty}{\longrightarrow} kJ_{k-1}(\beta-\alpha) + \alpha J_k(\beta-\alpha) \;=\; \beta J_k(\beta-\alpha) \;=\; \frac{1}{\gamma_1}\frac{k!}{(\gamma_1{}^{-1} - \gamma_2{}^{-1})^{k+1}} \;=\; c_k
\end{aligned}$$

## References


[1] J. Beirlant, G. Dierckx, A. Fils-Villetard, A. Guillou, Estimation of the extreme value index and extreme quantiles under random censoring, Extremes 10 (2007) 151–174.
[2] S. Benchaira, D. Meraghni, A.Necir, On the estimation of the extreme value index for randomly right-truncated data and application.
[3] N. Bingham, C. Goldie, J. Teugels, Regular variation, Cambridge University Press, 1987.
[4] J. Einmahl, A. Fils-Villetard, A. Guillou, Statistics of extremes under random censoring, Bernoulli 14 (2008) 207–227.
[5] L. Gardes, G. Stupfler, Estimating extreme quantiles under random truncation, TEST 24 (2015) 207–227.
[6] L. de Haan, A. Ferreira, Extreme Value Theory : an introduction, Springer Series in Operations Research and Financial Engineering, Springer, 2006.
[7] R. Smith, Estimating tails of probability distributions, Annals of Statistics 15 (3) (1987) 1174–120.
[8] E. Strzalkowska-Kominiak, W. Stute, On the probability of holes in truncated samples, Journal of Statistical Planning and Inference 140 (2010) 1519–1528.
[9] W.Stute, Almost sure representations of the product-limit estimator for truncated data, Annals of statistics 21 (1) (1993) 146–156.
[10] M. Woodroofe, Estimating a distribution function with truncated data, Annals of statistics 13 (1) (1985) 163–177.
[11] J. Worms, R. Worms, New estimators of the extreme value index under random right censoring, for heavy-tailed distributions, Extremes 17 (2014) 337–358.