Hill estimator and extreme quantile estimator for functionals of approximated stochastic processes

Jaakko Pere*[1], Benny Avelin[2], Valentin Garino[2], Pauliina Ilmonen[1], and Lauri Viitasaari[2]

[1] *Aalto University School of Science, Finland*
[2] *Uppsala University Department of Mathematics, Sweden*

July 29, 2023

Abstract: We study the effect of approximation errors in assessing the extreme behaviour of univariate functionals of random objects. We build our framework into a general setting where estimation of the extreme value index and extreme quantiles of the functional is based on some approximated value instead of the true one. As an example, we consider the effect of discretisation errors in computation of the norms of paths of stochastic processes. In particular, we quantify connections between the sample size n (the number of observed paths), the number of the discretisation points m, and the modulus of continuity function ϕ describing the path continuity of the underlying stochastic process. As an interesting example fitting into our framework, we consider processes of form $Y(t) = \mathcal{R}Z(t)$, where $\mathcal{R}$ is a heavy-tailed random variable and the increments of the process Z have lighter tails compared to $\mathcal{R}$.

Keywords: Regular variation, Extreme value theory, Functional data analysis, Hill estimator, Extreme quantile estimation

AMS subject classification: 62G32, 60G70

1. Introduction

Heavy-tailed phenomena arise naturally in multiple different application areas such as in telecommunications, finance and insurance [12]. In the present article, we analyse the effect of approximation errors in assessing the extreme behaviour of univariate functionals of random objects. That is, we estimate the extreme value index and extreme quantiles of the univariate functional based on approximated values.

*Corresponding author: jaakko.pere@aalto.fi

As an example, we consider functionals given by the p-norm of a stochastic process Y, i.e. $X = \|Y\|_p$ for $1 \leq p \leq \infty$. With $p = \infty$, this could for instance be used to measure worst losses of an insurance company over certain period of time, in which case extreme losses would mean that capital adequacy is compromised. In practice it might be that one does not observe the whole paths $Y_i(t), t \in [0, 1]$ but instead we observe simple processes

$$\hat{Y}_i = \sum_{j=0}^{m-1} \mathbb{1}_{[t_j, t_{j+1})} Y_i(t_j),$$

induced by the observed $n \times m$ discrete values $Y_i(t_j)$, where $0 = t_0 < t_1 < \cdots < t_{m-1} < t_m = T$. This leads to an approximation error when the norm $\|Y\|_p$ is evaluated by using values $\hat{Y}_i$. We quantify the connection between sample size n and the number of discretisation points m, depending on the regularity of the sample paths Y_i. Roughly speaking, heavier tails or lower path regularity of Y has to be compensated with larger m. We focus on the heavy tailed random objects, in which case it is natural to rely on the well-known Hill estimator, see [2, 12] for a review.

On a related literature on assessing extreme behaviour of stochastic processes, we mention [8, 9] concerning stochastic processes in a Skorokhod space and [10] in the context of functional principal component analysis (PCA). However, in those articles, the authors did not quantify the effect of approximation error which is our main focus here.

The rest of the article is organized as follows. In Section 2 we give the notation and rate for the approximation error in order to have the standard asymptotic results for the Hill estimator and the extreme quantile estimator, when approximations of regularly varying random variables are used in the estimation. In Section 3 we illustrate our results by considering the case where Y_i are random elements of the space $L^p[0, 1]$, $p \in [1, \infty]$.

2. General framework

Let $X_1, \ldots, X_n$ be i.i.d. random variables with a common distribution F. The tail quantile function U corresponding the distribution F is defined by

$$U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right), \quad t > 1,$$

where $f^{\leftarrow}$ denotes the left-continuous inverse of a nondecreasing function f. Throughout the article, if needed, we stress that the tail quantile function U corresponds to a certain random variable X by denoting U_X. Order statistics corresponding to the sample $X_1, \ldots, X_n$ are denoted by $X_{1,n} \leq \cdots \leq X_{n,n}$.

In this section, we consider a general approximations $\hat{X}_i$ of "true" random variables X_i that are i.i.d. and heavy-tailed. In our application into functional data on Section 3, we set $X_i = h(Y_i)$, where Y_i is some stochastic process on the interval $[0, T]$ and h is some real-valued map.

Definition 1 (Heavy-tailed random variable)**.** *Random variable X is heavy-tailed with extreme value index $\gamma > 0$ and we denote $X \in RV_\gamma$ if the corresponding tail quantile function U is regularly varying with the same index $\gamma > 0$. That is,*

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\gamma,$$

for all $x > 0$.

Let $k \in \{1, \dots, n-1\}$. We define the Hill estimator and the extreme quantile estimator for $x_p = F^{\leftarrow}(1-p) = U(1/p)$ as

$$\tilde{\gamma}_n = \frac{1}{k}\sum_{i=0}^{k-1} \log\left(\frac{X_{n-i,n}}{X_{n-k,n}}\right) \quad \text{and} \quad \tilde{x}_p = X_{n-k,n}\left(\frac{k}{np}\right)^{\tilde{\gamma}_n}, \tag{1}$$

where $p \in (0,1)$ is small. The Hill estimator $\tilde{\gamma}_n$, introduced in [7], is one of the most well-known estimators for the extreme value index $\gamma > 0$ in the case of regularly varying random variables. Consistency of the Hill estimator was first established in [11] under i.i.d. heavy-tailed random variables, when $k = k_n \to \infty$, $k/n \to 0$, as $n \to \infty$. Also, asymptotic normality of the Hill estimator has been studied by various authors [4, 5]. In asymptotic results for the extreme quantile estimator, it is often assumed that $p = p_n$ and $p \to 0$ fast as $n \to \infty$. This is necessary since otherwise quantile x_p would not be extreme for a sufficiently large sample size n. For a review about the Hill estimator and the extreme quantile estimator we refer to [2].

For the asymptotic normality of the Hill estimator $\tilde{\gamma}_n$ and the extreme quantile estimator $\tilde{x}_p$, the regular variation condition is not sufficient, but instead a second-order condition is required, see e.g. [2, Theorem 3.2.5] and [2, Theorem 4.3.8].

Definition 2 (Second-order regular variation)**.** *We say that X satisfies second-order regular variation condition and denote $X \in 2RV_{\gamma,\rho}$ if the corresponding tail quantile function U satisfies*

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}$$

for all $x > 0$, where $\gamma > 0$, $\rho \le 0$ and A is a positive or negative function with $\lim_{t\to\infty} A(t) = 0$. *Here for $\rho = 0$, the right-hand side is interpreted as $x^\gamma \log x$.*

In our setting, we do not observe $X_1, \dots, X_n$ directly but instead, we have approximations $\hat{X}_1, \dots, \hat{X}_n$ that are used in the estimation. We define the Hill estimator and the extreme quantile estimator computed with the approximated values as

$$\hat{\gamma}_n = \frac{1}{k}\sum_{i=0}^{k-1} \log\left(\frac{\hat{X}_{n-i,n}}{\hat{X}_{n-k,n}}\right) \quad \text{and} \quad \hat{x}_p = \hat{X}_{n-k,n}\left(\frac{k}{np}\right)^{\hat{\gamma}_n}. \tag{2}$$

The following result gives insights how the approximation error $\max_i |\hat{X}_i - X_i|$ has to behave in order to have the standard asymptotic results for the Hill estimator and the extreme quantile estimator, when approximated values $\hat{X}_i$ are used in the estimation.

Theorem 1. *Let $X_1, \ldots, X_n$ be i.i.d. copies of a heavy-tailed random variable $X \in RV_\gamma$ with extreme value index $\gamma > 0$. Assume that $k = k_n \to \infty$, $k/n \to 0$, as $n \to \infty$. Let $\hat{X}_i, \ldots, \hat{X}_n$ be a sequence of random variables and denote*

$$C_n = \frac{\max_i |\hat{X}_i - X_i|}{U(n/k)}.$$

Let $\hat{\gamma}_n$ and $\hat{x}_p$ be as in Equation (2).

1. *We have*
$$\hat{\gamma}_n \xrightarrow{\mathbb{P}} \gamma, \quad \text{as } n \to \infty,$$
provided that $C_n = o_{\mathbb{P}}(1)$.

2. *Suppose further that $X \in 2RV_{\gamma,\rho}$ with $\gamma > 0$ and $\rho \leq 0$, and that $\lim_{n\to\infty} \sqrt{k}A(n/k) = \lambda \in \mathbb{R}$, where A is the auxiliary function from Definition 2. Let*
$$\Gamma \sim N\left(\frac{\lambda}{1-\rho}, \gamma^2\right). \tag{3}$$
Then
$$\sqrt{k}(\hat{\gamma}_n - \gamma) \xrightarrow{d} \Gamma,$$
provided that $\sqrt{k}C_n = o_{\mathbb{P}}(1)$.

3. *Furthermore, assume that $X \in 2RV_{\gamma,\rho}$ with $\gamma > 0$ and $\rho < 0$ and that $\lim_{n\to\infty} \sqrt{k}A(n/k) = \lambda \in \mathbb{R}$, where A is the auxiliary function from Definition 2. Assume also that $p = p_n$, $np = o(k)$ and $\log(np) = o(\sqrt{k})$, as $n \to \infty$, and denote $d_n = k/(np)$. Then*
$$\frac{\sqrt{k}}{\log d_n}\left(\frac{\hat{x}_p}{x_p} - 1\right) \xrightarrow{d} \Gamma,$$
provided that $\sqrt{k}C_n = o_{\mathbb{P}}(1)$, where Γ is given by (3).

Below remark shows that the rates for approximation error given in above theorem can be written without the function U by using the fact that a regularly varying function $f \in RV_\alpha$, $\alpha \in \mathbb{R}$, can be represented as $f(x) = L(x)x^\alpha$, where L is a slowly varying function. Additionally, under second-order regular variation with parameters $\gamma > 0$ and $\rho < 0$ we have even simpler expressions for the rates.

Remark 1. *Note that the tail quantile function U of X can be written as $U(x) = L(x)x^\gamma$, where $L(x)$ is a slowly varying function, i.e. $L \in RV_0$. It follows from Karamata's representation theorem [2, Theorem B.1.6] that for $\varepsilon > 0$ we have $x^{-\varepsilon}L(x) \to 0$ and $x^\varepsilon L(x) \to \infty$, as $n \to \infty$, and consequently*

$$C_n = o_{\mathbb{P}}\left(\left(\frac{k}{n}\right)^{\gamma-\varepsilon} \max_i |X_i - \hat{X}_i|\right)$$

for any $\varepsilon > 0$. Similarly, if the second order condition $X \in 2RV_{\gamma,\rho}$ holds with $\gamma > 0$ and $\rho < 0$ then $U(t)/t^\gamma \to c \in (0, \infty)$, as $t \to \infty$, see [2, page 49]. That is, the tail quantile function U can be represented as $U(t) = L(t)t^\gamma$, where $L(t) \to c$, as $t \to \infty$. In this case

$$C_n = O_{\mathbb{P}}\left(\left(\frac{k}{n}\right)^\gamma \max_i |X_i - \hat{X}_i|\right).$$

3. Regularly varying L^p-norms

In this section we assume that $Y_1, \ldots, Y_n$ are random elements of a Banach space $L^p[0,1]$ with the norm

$$\|y\|_p = \begin{cases} \left(\int_0^1 |y(t)|^p \, \mathrm{d}t\right)^{\frac{1}{p}}, & p \in [1, \infty), \\ \sup_{t\in[0,1]} |y(t)|, & p = \infty, \end{cases}$$

and we are interested in the extreme behavior of the norms $X_i = h(Y_i)$ with $h(y) = \|y\|_p$. Note that other choices of the functional $h : L^p[0,1] \mapsto \mathbb{R}$ are possible as long as h has suitable continuity properties, but for the simplicity of our presentation we restrict ourselves to the case of the norms. Now in order to obtain our approximation, in practice it is often the case that our stochastic processes Y_i are observed only at discrete times j/m for $j \in \{0, \ldots, m-1\}$, leading to a discretisation error of the integral. We consider the natural approximations

$$\hat{X}_i = \begin{cases} \left(\frac{1}{m}\sum_{j=0}^{m-1} \left|Y_i\left(\frac{j}{m}\right)\right|^p\right)^{\frac{1}{p}}, & p \in [1, \infty) \\ \max_{j\in\{0,\ldots,m-1\}} \left|Y_i\left(\frac{j}{m}\right)\right|, & p = \infty \end{cases}. \tag{4}$$

Then, with rather standard regularity assumptions, we can show that this setup falls into the framework of the result of Theorem 1.

Theorem 2. *Let $Y_1, \ldots, Y_n$ be i.i.d. random elements of a Banach space $L^p[0,1]$, $1 \leq p \leq \infty$, and let $X_i = \|Y_i\|_p$. Suppose $X = \|Y\|_p \in RV_\gamma$ with extreme value index $\gamma > 0$, and let $\hat{X}_1, \ldots, \hat{X}_n$ be as in Equation (4). Suppose also the continuity criterion*

$$|Y_i(t) - Y_i(s)| \leq V_i \phi(|t - s|),$$

where V_i are i.i.d copies of a heavy-tailed random variable $V \in RV_{\gamma'}$ with extreme value index $\gamma' \geq \gamma$, and the modulus of continuity $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a continuous nondecreasing function with $\phi(\delta) \to 0$, as $\delta \to 0$. Then

$$\max_i |X_i - \hat{X}_i| = O_{\mathbb{P}}\left(\phi\left(\frac{1}{m}\right) U_V(n)\right).$$

By combining Theorem 1 and Theorem 2 leads to the following result that allows to study extreme behaviour of the norms.

Corollary 1. *Suppose that assumptions of Theorem 2 hold. Furthermore, assume that $k = k_n \to \infty$, $k/n \to 0$, as $n \to \infty$ and let $\hat{\gamma}_n$ and $\hat{x}_p$ be as in Equation (2).*

1. *We have*

$$\hat{\gamma}_n \xrightarrow{\mathbb{P}} \gamma, \quad \text{as } n \to \infty,$$

provided that

$$\phi\left(\frac{1}{m}\right)\frac{U_V(n)}{U_X(n/k)} \to 0, \quad \text{as } n \to \infty.$$

2. *Suppose further that* $X \in 2RV_{\gamma,\rho}$ *with* $\gamma > 0$ *and* $\rho \leq 0$*, and that* $\lim_{n\to\infty} \sqrt{k}A(n/k) = \lambda \in \mathbb{R}$*, where* A *is the auxiliary function from Definition 2. Then for* Γ *given by* (3)*, we have*

$$\sqrt{k}(\hat{\gamma}_n - \gamma) \xrightarrow{d} \Gamma,$$

provided that

$$\sqrt{k}\phi\left(\frac{1}{m}\right)\frac{U_V(n)}{U_X(n/k)} \to 0, \quad \text{as } n \to \infty.$$

3. *Furthermore, assume that* $X \in 2RV_{\gamma,\rho}$ *with* $\gamma > 0$ *and* $\rho < 0$ *and that* $\lim_{n\to\infty} \sqrt{k}A(n/k) = \lambda \in \mathbb{R}$*, where* A *is the auxiliary function from Definition 2. Assume that* $p = p_n$*,* $np = o(k)$ *and* $\log(np) = o(\sqrt{k})$*, as* $n \to \infty$*. Denote* $d_n = k/(np)$*. Then for* Γ *given by* (3)*, we have*

$$\frac{\sqrt{k}}{\log d_n}\left(\frac{\hat{x}_p}{x_p} - 1\right) \xrightarrow{d} \Gamma,$$

provided that

$$\sqrt{k}\phi\left(\frac{1}{m}\right)\frac{U_V(n)}{U_X(n/k)} \to 0, \quad \text{as } n \to \infty.$$

Below remark states that with similar line of thought as in Remark 1, conditions for the approximation error can be written without the tail quantile functions U_X and U_V.

Remark 2. *Suppose that assumptions of Theorem 2 hold. Then* $U_X(x) = L_X(x)x^{\gamma}$ *and* $U_V(x) = L_V(x)x^{\gamma'}$*, where* L_X *and* L_V *are slowly varying functions, and thus,*

$$\phi\left(\frac{1}{m}\right)\frac{U_V(n)}{U_X(n/k)} = O\left(\phi\left(\frac{1}{m}\right)k^{\gamma-\varepsilon}n^{\gamma'-\gamma+\varepsilon+\varepsilon'}\right).$$

for any $\varepsilon, \varepsilon' > 0$*. Additionally, if* $X \in 2RV_{\gamma,\rho}$ *and* $V \in 2RV_{\gamma',\rho'}$*, where* $\gamma' \geq \gamma \geq 0$ *and* $\rho, \rho' < 0$*, then*

$$\phi\left(\frac{1}{m}\right)\frac{U_V(n)}{U_X(n/k)} = O\left(\phi\left(\frac{1}{m}\right)k^{\gamma}n^{\gamma'-\gamma}\right).$$

This shows that, depending on the modulus of continuity ϕ*, one needs larger* m *in order to compensate effects arising from the chosen intermediate sequence* k *and different indices* γ', γ *for* V *and* X*.*

A natural class of suitable stochastic processes satisfying all the assumptions of Theorem 2 involves products $Y(t) = RZ(t)$, where R is a heavy-tailed random variable and Z is a continuous process having lighter tails than R.

Theorem 3. *Let $\mathcal{R} \in RV_\gamma$ be heavy-tailed positive random variable with an extreme value index $\gamma > 0$, and let Z be a stochastic process such that $Z(0) = 0$ almost surely and, for some $\varepsilon > 0$,*

$$\mathbb{E}\left[|Z(t) - Z(s)|^{\kappa_\varepsilon}\right] \leq K\,|t-s|^{1+\beta} \quad \forall\, t, s \in [0,1], \tag{5}$$

where $\kappa_\varepsilon = 1/\gamma + \varepsilon$. Suppose that $\mathcal{R}$ and Z are independent and set

$$Y(t) = \mathcal{R}Z(t), \quad t \in [0,1]. \tag{6}$$

Then

1. *$X = \|Y\|_p \in RV_\gamma$ for any $1 \leq p \leq \infty$,*
2. *$|Y(t) - Y(s)| \leq V|t-s|^\eta$ for all $\eta \in (0, \beta/\kappa_\varepsilon)$, where $V \in RV_\gamma$.*

Note that condition (5) is a Kolmogorov continuity criterion, where we have assumed that Z has more moments than $1/\gamma$. Let us shed some light how the approximation error for the family of stochastic processes defined in Equation (6) has to behave in order to have standard asymptotic results for the Hill estimator $\hat{\gamma}_n$ and for the extreme quantile estimator $\hat{x}_p$. Let $k = \lfloor n^\lambda \rfloor$, where $\lambda \in (0,1)$. Then in view of Remark 2, for the consistency of the Hill estimator $\hat{\gamma}_n$ it is sufficient to have

$$\overbrace{\left(\frac{1}{m}\right)^{(\beta/\kappa_\varepsilon)-\varepsilon'}}^{\text{Term I}} \underbrace{n^{\lambda\gamma+\varepsilon'}}_{\text{Term II}} \to 0, \quad \text{as } n \to \infty, \tag{7}$$

for any $\varepsilon' > 0$, where β and ε are constants from Theorem 3. We can make a few observations by investigating Relation (7). Firstly, as discretization level m increases Term I decays faster, and similarly, by choosing a smaller k (that is, smaller λ in this case) Term II grows slower. Consequently, heavy-tailedness and low regularity can be compensated by choosing smaller k and larger discretization level m, respectively. Secondly, there is a connection between regularity of the process defined in Equation (6) and the heavy-tailedness of the norm $\|Y\|_p$. That is, as γ increases regularity increases and Term I decays to zero faster, but then also Term II grows faster. An interesting example is the case where the process $Z(t)$ is an H-Hölder continuous Gaussian process (or more generally, H-Hölder continuous hypercontractive process). Then the increments have all the moments satisfying

$$\mathbb{E}|Z(t) - Z(s)|^p \leq K_p|t-s|^{Hp}.$$

This would yield condition

$$\left(\frac{1}{m}\right)^{H-\varepsilon'} n^{\lambda\gamma+\varepsilon'} \to 0, \quad \text{as } n \to \infty, \tag{8}$$

for any $\varepsilon' > 0$.

Acknowledgements: Jaakko Pere gratefully acknowledges support from the Vilho, Yrjö and Kalle Väisälä Foundation. Pauliina Ilmonen gratefully acknowledges support from the Academy of Finland via the Centre of Excellence in Randomness and Structures, decision number 346308.

References

[1] L. Breiman. On some limit theorems similar to the arc-sin law. *Theory of Probability & Its Applications*, 10(2):323–331, 1965.

[2] L. De Haan and A. Ferreira. Extreme value theory: An introduction, 2007.

[3] A. M. Garsia, E. Rodemich, H. Rumsey, and M. Rosenblatt. A real variable lemma and the continuity of paths of some gaussian processes. *Indiana University Mathematics Journal*, 20(6):565–578, 1970.

[4] E. Haeusler and J. L. Teugels. On asymptotic normality of Hill's estimator for the exponent of regular variation. *The Annals of Statistics*, pages 743–756, 1985.

[5] P. Hall. On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(1):37–42, 1982.

[6] M. Heikkilä, Y. Dominicy, and P. Ilmonen. On multivariate separating Hill estimator under estimated location and scatter. *Statistics*, 53(2):301–320, 2019.

[7] B. M. Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.

[8] H. Hult and F. Lindskog. Extremal behavior of regularly varying stochastic processes. *Stochastic Processes and their applications*, 115(2):249–274, 2005.

[9] H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publications de l'Institut Mathématique*, 80(94):121–140, 2006.

[10] M. Kim and P. Kokoszka. Hill estimator of projections of functional data on principal components. *Statistics*, 53(4):699–720, 2019.

[11] D. M. Mason. Laws of large numbers for sums of extreme values. *The Annals of Probability*, pages 754–764, 1982.

[12] S. I. Resnick. *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media, 2007.

Appendices

A. Proofs of Section 2

Before proving the main result of Section 2 we need the following lemma.

Lemma 1. *Let $X_1, \ldots, X_n$ be i.i.d. copies of a heavy-tailed random variable $X \in RV_\gamma$ with extreme value index $\gamma > 0$. Assume that $k = k_n \to \infty$, $k/n \to 0$, as $n \to \infty$. Let $\hat{X}_1, \ldots, \hat{X}_n$ be a sequence of random variables. Then there exists a sequence of positive random variables K_n such that*

$$|\hat{X}_{\ell,n} - X_{\ell,n}| \leq K_n X_{\ell,n}, \quad \text{for} \quad \ell \in \{n-k, \ldots, n\}$$

and

$$K_n = O_{\mathbb{P}}\left(\frac{\max_i |\hat{X}_i - X_i|}{U(n/k)}\right).$$

Proof of Lemma 1. In the proof we use the shorthand notation $[n] = \{1, \ldots, n\}$. Denote $\tilde{C}_n = \max_i |\hat{X}_i - X_i|$. For all $m \in [n]$ we have

$$|\hat{X}_m - X_m| \leq \tilde{C}_n \iff X_m - \tilde{C}_n \leq \hat{X}_m \leq X_m + \tilde{C}_n.$$

Denote sets of indices by

$$I_\ell = \{i \in [n] : X_i \leq X_{\ell,n}\} \quad \text{and}$$
$$J_\ell = \{j \in [n] : X_j \geq X_{\ell,n}\},$$

where $\ell \in \{n-k, \ldots, n\}$. Since function $f(x) = x + c$, $c \in \mathbb{R}$ is increasing we have

$$\begin{cases} \hat{X}_i \leq X_{\ell,n} + \tilde{C}_n & \forall i \in I_\ell \\ \hat{X}_j \geq X_{\ell,n} - \tilde{C}_n & \forall j \in J_\ell \end{cases}.$$

Then, since $\# I_\ell = \ell$ and $\# J_\ell = n - \ell + 1$, it follows that

$$\begin{cases} \hat{X}_{\ell,n} \leq X_{\ell,n} + \tilde{C}_n \\ \hat{X}_{\ell,n} \geq X_{\ell,n} - \tilde{C}_n \end{cases},$$

and thus

$$|\hat{X}_{\ell,n} - X_{\ell,n}| \leq \tilde{C}_n.$$

By combining [2, Theorem 1.2.1] and [2, Corollary 1.2.10] it follows that right endpoint $\sup\{x \in \mathbb{R} : F(x) < 1\}$ of a heavy-tailed distribution is infinite. Then by [6, Lemma 1] we have

$$X_{l,n} \overset{a.s}{\to} \infty \quad \text{as } n \to \infty.$$

Consequently, almost surely and for large enough n we have

$$|X_{\ell,n}| = X_{\ell,n}.$$

This leads to

$$|X_{l,n}| \frac{|\hat{X}_{l,n} - X_{l,n}|}{|X_{l,n}|} \leq X_{l,n} \frac{\tilde{C}_n}{X_{l,n}} \leq X_{l,n} \frac{\tilde{C}_n}{X_{n-k,n}} = X_{l,n} \frac{\tilde{C}_n}{U(n/k)} \frac{U(n/k)}{X_{n-k,n}}.$$

Hence we may set

$$K_n = \frac{\tilde{C}_n}{U(n/k)} \frac{U(n/k)}{X_{k,n}},$$

where (see Step 1 in the proof of [12, Theorem 4.2] for details)

$$\frac{U(n/k)}{X_{n-k}} \xrightarrow{\mathbb{P}} 1, \quad \text{as } n \to \infty.$$

This gives

$$K_n = O_{\mathbb{P}}\left(\frac{\tilde{C}_n}{U(n/k)}\right)$$

completing the proof. □

Now we are ready to prove Theorem 1.

Proof of Theorem 1. Let $\tilde{\gamma}_n$ and $\tilde{x}_p$ be as in Equation (1). By applying triangle inequality two times and using the fact that $|\log(a/b)| = |\log(b/a)|$ for $a, b \neq 0$ we obtain

$$|\tilde{\gamma}_n - \hat{\gamma}_n| \leq \frac{1}{k}\sum_{i=0}^{k-1}\left|\log\left(\frac{\hat{X}_{n-i,n}}{X_{n-i,n}}\right)\right| + \left|\log\left(\frac{\hat{X}_{n-k,n}}{X_{n-k,n}}\right)\right|.$$

Let $\ell \in \{n-k, \ldots, n\}$. Next we prove that if

$$0 < \frac{|X_{\ell,n} - \hat{X}_{\ell,n}|}{X_{\ell,n}} < 1, \tag{9}$$

then

$$\left|\log\left(\frac{\hat{X}_{\ell,n}}{X_{\ell,n}}\right)\right| \leq \left|\log\left(1 - \frac{|X_{\ell,n} - \hat{X}_{\ell,n}|}{X_{\ell,n}}\right)\right|. \tag{10}$$

If $X_{\ell,n} \geq \hat{X}_{\ell,n}$ then Relation (10) holds with equality. Thus consider the case $X_{\ell,n} < \hat{X}_{\ell,n}$. In this case Relation (10) becomes

$$\log\left(\frac{\hat{X}_{\ell,n}}{X_{\ell,n}}\right) \leq \log\left(\left(2 - \frac{\hat{X}_{\ell,n}}{X_{\ell,n}}\right)^{-1}\right),$$

and Condition (9) becomes

$$1 < \frac{\hat{X}_{\ell,n}}{X_{\ell,n}} < 2,$$

from which (10) follows from the fact

$$x \leq (2-x)^{-1}, \quad \text{for} \quad 1 < x < 2.$$

In all three parts of the theorem, we assume $C_n = o_{\mathbb{P}}(1)$. Consequently, $K_n = o_{\mathbb{P}}(1)$, where K_n is the sequence from Lemma 1. Then Condition (9) holds almost surely for large enough n and

$$\left|\log\left(\frac{\hat{X}_{\ell,n}}{X_{\ell,n}}\right)\right| \leq |\log(1-K_n)|.$$

This gives

$$|\tilde{\gamma}_n - \hat{\gamma}_n| \leq 2|\log(1-K_n)| \tag{11}$$

and assuming $K_n \leq 1/2$ (that is valid for large enough n since $K_n = o_{\mathbb{P}}(1)$) we get

$$\begin{aligned} |\log(1-K_n)| &= -\log(1-K_n) = \sum_{m=1}^{\infty} \frac{K_n^m}{m} \\ &= K_n \sum_{m=1}^{\infty} \frac{K_n^{m-1}}{m} \leq K_n \sum_{m=1}^{\infty} \frac{1}{2^{m-1}m} = K_n \log(4). \end{aligned} \tag{12}$$

Combining (11) and (12) now yields

$$|\tilde{\gamma}_n - \hat{\gamma}_n| = O_{\mathbb{P}}(C_n). \tag{13}$$

Now the consistency result of Part 1 follows immediately by combining Relation (13) and [2, Theorem 3.2.2]. Similarly, asymptotic normality of the Hill estimator stated in Part 2 follows immediately by combining Relation (13) and Theorem 3.2.5 from [2, Theorem 3.2.5]. Hence it remains to prove the asymptotic normality of the extreme quantile estimator stated in Part 3. For this we have

$$\frac{\sqrt{k}}{\log d_n}\left(\frac{\hat{x}_p}{x_p} - 1\right) = \underbrace{\frac{\sqrt{k}}{\log d_n}\left(\frac{\tilde{x}_p}{x_p} - 1\right)}_{\mathrm{I}} + \underbrace{\frac{\sqrt{k}}{\log d_n}\left(\frac{\hat{x}_p}{x_p} - \frac{\tilde{x}_p}{x_p}\right)}_{\mathrm{II}},$$

where term I converges to Γ in distribution by [2, Theorem 4.3.8]. Thus, it is sufficient to prove that term II converges to zero in probability. We have already showed that under the assumptions of Part 3 we have $\sqrt{k}(\hat{\gamma}_n - \gamma) = O_{\mathbb{P}}(1)$. Combining this to $\log(np) = o\left(\sqrt{k}\right)$ gives

$$d_n^{\hat{\gamma}_n - \gamma} = \exp\left(\sqrt{k}(\hat{\gamma}_n - \gamma)\left(\frac{\log(k)}{\sqrt{k}} - \frac{\log(np)}{\sqrt{k}}\right)\right) \xrightarrow{\mathbb{P}} 1, \quad n \to \infty. \tag{14}$$

Similar argument gives also that $d_n^{\tilde{\gamma}_n - \gamma} \xrightarrow{\mathbb{P}} 1$, as $n \to \infty$. Moreover, since $X_{n-k,n} \xrightarrow{a.s.} \infty$ as $n \to \infty$ the quantity $\hat{X}_{n-k,n}/X_{n-k,n}$ is eventually well-defined, and by combining Lemma 1 with $K_n = O_{\mathbb{P}}(C_n)$ shows that

$$\frac{\hat{X}_{n-k_n,n}}{X_{n-k_n,n}} = 1 + O_{\mathbb{P}}(C_n). \tag{15}$$

Combining (14) and (15) now yields

$$\hat{x}_p - \tilde{x}_p = \left(\frac{\hat{X}_{n-k,n}}{X_{n-k,n}} d_n^{\hat{\gamma}_n - \gamma} d_n^{\gamma - \tilde{\gamma}_n} - 1\right)\tilde{x}_p = O_{\mathbb{P}}(C_n)\tilde{x}_p.$$

Thus we have

$$\mathrm{II} = \frac{\sqrt{k}}{\log d_n} O_{\mathbb{P}}(C_n) \frac{\tilde{x}_p}{x_p} \xrightarrow{\mathbb{P}} 0, \quad n \to \infty,$$

since $d_n \to \infty$, $\tilde{x}_p/x_p = 1 + o_{\mathbb{P}}(1)$ and $\sqrt{k}C_n = o_{\mathbb{P}}(1)$. This completes the proof. $\square$

B. Proofs of Section 3

Proof of Theorem 2. Let first $p \in [1, \infty)$ and define a simple process $\hat{Y}_i$ corresponding to discretization by

$$\hat{Y}_i = \sum_{j=0}^{m-1} \mathbb{1}_{[j/m,(j+1)/m)} Y_i\left(\frac{j}{m}\right).$$

Since intervals $[j/m, (j+1)/m)$ are disjoint,

$$\begin{aligned}
\|\hat{Y}_i\|_p &= \left(\int_0^1 |\hat{Y}_i(y)|^p \,\mathrm{d}t\right)^{1/p} = \left(\sum_{j=0}^{m-1} \int_{j/m}^{(j+1)/m} \left|Y_i\left(\frac{j}{m}\right)\right|^p \,\mathrm{d}t\right)^{1/p} \\
&= \left(\frac{1}{m}\sum_{j=0}^{m-1} \left|Y_i\left(\frac{j}{m}\right)\right|^p\right)^{1/p} = \hat{X}_i.
\end{aligned}$$

Minkowski inequality leads to

$$\begin{aligned}
|X_i - \hat{X}_i| &= \left|\|Y_i\|_p - \left\|\hat{Y}_i\right\|_p\right| \le \left\|Y_i - \hat{Y}_i\right\|_p \\
&= \left(\sum_{j=0}^{m-1} \int_{j/m}^{(j+1)/m} |Y_i(t) - Y_i(j/m)|^p \,\mathrm{d}t\right)^{1/p} \le V_i\phi(1/m),
\end{aligned}$$

and consequently,

$$\max_{i\in\{1,\dots,n\}} |X_i - \hat{X}_i| \le \phi\left(\frac{1}{m}\right) \max_{i\in\{1,\dots,n\}} V_i. \tag{16}$$

Clearly now the random variables i.i.d. and $V_i \in RV_\gamma$ for $\gamma > 0$, which is equivalent to the fact that

$$\frac{\max_i V_i}{U_V(n)} \xrightarrow{d} M, \tag{17}$$

where M is random variable with $\mathbb{P}(M \le x) = \exp\left(-x^{-1/\gamma}\right)$ (see [2, Corollary 1.2.4] and [2, Corollary 1.2.10] for details). Now the result follows by combining Relations (16) and (17). Finally, the case $p = \infty$ follows with similar arguments by observing that

$$\left|\sup_{t\in[0,1]} |Y_i(t)| - \max_{1\le j\le m} \left|Y_i\left(\frac{j}{m}\right)\right|\right| \le V_i\phi(1/m).$$

$\square$

The main technical result required for the proof of Theorem 3 is the following regularity result due to Garsia, Rodemich, Rumsay [3].

Lemma 2. *Let $\theta, \alpha > 0$ such that $\alpha\theta > 1$ and let f be a continuous function. Then, $\forall s, t \in [0, 1]$,*

$$|f(t) - f(s)| \le K_{\alpha,\theta}|t - s|^{\alpha - \frac{1}{\theta}} \left(\int_0^1 \int_0^1 \frac{|f(u) - f(v)|^\theta}{|u - v|^{\alpha\theta+1}} \,\mathrm{d}u\,\mathrm{d}v\right)^{\frac{1}{\theta}}.$$

We will also exploit the following lemma, taken from [1], stating that the product of independent heavy-tailed and light-tailed random variables is heavy-tailed.

Lemma 3. *Let X and Z be nonnegative independent random variables. Assume that $X \in RV_\gamma$ with $\gamma > 0$ and $\mathbb{E}Z^{\frac{1}{\gamma}+\varepsilon} < \infty$ for some $\varepsilon > 0$. Then XZ is regularly varying with index γ.*

Now we are ready to prove Theorem 3.

Proof of Theorem 3. Let $\delta \in (0,1)$, denote $\alpha = ((1-\delta)\beta + 1)/\kappa_\varepsilon$ and set $\eta' = \alpha - 1/\kappa_\varepsilon$. Since $\eta' \in (0, \beta/\kappa_\varepsilon)$, we have by Kolmogorov continuity theorem that

$$|Z(t) - Z(s)| \leq M|t-s|^{\eta'},$$

where

$$M = \sup_{0 \leq t < s \leq 1} \frac{|Z(t) - Z(s)|}{|t-s|^{\eta}}.$$

On the other hand, since $\alpha\kappa_\varepsilon > 1$, by Lemma 2 we get

$$M \leq K_{\alpha,\kappa_\varepsilon} \left(\int_0^1 \int_0^1 \frac{|Z(u) - Z(v)|^{\kappa_\varepsilon}}{|u-v|^{\alpha\kappa_\varepsilon + 1}} \, \mathrm{d}u \, \mathrm{d}v \right)^{1/\kappa_\varepsilon}.$$

Now, by combining Assumption (5) and Fubini's theorem we have

$$\begin{aligned} \mathbb{E}M^{\kappa_\varepsilon} &= O\left(\int_0^1 \int_0^1 \frac{\mathbb{E}\left[|Z(u) - Z(v)|^{\kappa_\varepsilon}\right]}{|u-v|^{\alpha\kappa_\varepsilon+1}} \, \mathrm{d}u \, \mathrm{d}v \right) \\ &= O\left(\int_0^1 \int_0^1 \frac{|u-v|^{1+\beta}}{|u-v|^{\alpha\kappa_\varepsilon+1}} \, \mathrm{d}u \, \mathrm{d}v \right) < \infty, \end{aligned}$$

since $(1+\beta) - (\alpha\kappa_\varepsilon + 1) = \delta\beta - 1 > -1$. As $\delta \in (0,1)$ is arbitrary, we get

$$|Z(t) - Z(s)| \leq M|t-s|^{\eta} \quad \forall \eta \in (0, \beta/\kappa_\varepsilon), \tag{18}$$

where $\mathbb{E}M^{\kappa_\varepsilon} < \infty$. With this preliminary step, we are now ready to prove Claims 1 and 2 of Theorem 3.

Proof of Claim 1. We only consider the case $p \in [1, \infty)$ since again the case $p = \infty$ can be handled similarly. By Equation (18) and the fact $Z(0) = 0$ we have

$$\|Z\|_p^{\kappa_\varepsilon} = \left(\int_0^1 |Z(t) - Z(0)|^p \, \mathrm{d}t \right)^{\kappa_\varepsilon/p} \leq M^{\kappa_\varepsilon} \underbrace{\left(\int_0^1 t^{\eta p} \, \mathrm{d}t \right)^{\kappa_\varepsilon/p}}_{=c<\infty}.$$

Thus,

$$\mathbb{E}\|Z\|_p^{\kappa_\varepsilon} \leq c\mathbb{E}M^{\kappa_\varepsilon} \leq \infty.$$

By assumption, Z and $\mathcal{R}$ are mutually independent and hence $\mathcal{R}$ and $\|Z\|_p$ are independent as well. It follows then by Lemma 3 that $\|Y\|_p = \mathcal{R}\|Z\|_p \in RV_\gamma$.

Proof of Claim 2. By Equation (18) we have

$$|Y(t) - Y(s)| \leq V|t-s|^{\eta} \quad \forall \, \eta \in (0, \beta/\kappa_\varepsilon),$$

where $V = \mathcal{R}M$. Here M is a measurable map of Z that is independent of $\mathcal{R}$, so then $\mathcal{R}$ and M are independent as well. Thus, again by Lemma 3, we have $V_i \in RV_\gamma$. $\square$