

# *Estimation of the finite right endpoint in the Gumbel domain*

Article

Accepted Version

Fraga Alves, I. and Neves, C. ORCID: <https://orcid.org/0000-0003-1201-5720> (2014) Estimation of the finite right endpoint in the Gumbel domain. *Statistica Sinica*, 24. pp. 1811-1835. ISSN 1017-0405 doi: <https://doi.org/10.5705/ss.2013.183>  
Available at <https://centaur.reading.ac.uk/67416/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://www3.stat.sinica.edu.tw/statistica/oldpdf/A24n417.pdf>

To link to this article DOI: <http://dx.doi.org/10.5705/ss.2013.183>

Publisher: Academia Sinica, Institute of Statistical Science

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Estimation of the finite right endpoint in the Gumbel domain

Isabel Fraga Alves

Cláudia Neves

CEAUL and DEIO

CEAUL and University of Aveiro

FCUL, University of Lisbon

## Abstract

A simple estimator for the finite right endpoint of a distribution function in the Gumbel max-domain of attraction is proposed. Large sample properties such as consistency and the asymptotic distribution are derived. A simulation study is also presented.

## 1 Introduction

Let  $X_{n,n} \geq X_{n-1,n} \geq \dots \geq X_{1,n}$  be the order statistics from the sample  $X_1, X_2, \dots, X_n$  of i.i.d. random variables with common (unknown) distribution function  $F$ . Let  $x^F$  denote the right endpoint of  $F$ . We shall assume that the distribution function  $F$  has a finite right endpoint, i.e.  $x^F := \sup\{x : F(x) < 1\} \in \mathbb{R}$ .

The fundamental result for extreme value theory is due in various degrees of generality to Fisher and Tippett (1928), Gnedenko (1943), de Haan (1970) and Balkema and de Haan (1974). The extreme value theorem (or extremal types theorem) surprisingly restricts the class of all possible limiting distribution functions to only three different types, while the induced domains of attraction embrace a great variety of distribution functions. This is particularly true in the case of the Gumbel domain of attraction. In other words, if there exist constants  $a_n > 0$ ,  $b_n \in \mathbb{R}$  such that

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x), \tag{1}$$

for all  $x$ ,  $G$  non-degenerate, then  $G$  must be only one of the following:

$$\begin{aligned}\Psi_\alpha(x) &= \exp\{-(-x)^\alpha\}, \quad x < 0, \quad \alpha > 0, \\ \Lambda(x) &= \exp\{-\exp(-x)\}, \quad x \in \mathbb{R}, \\ \Phi_\alpha(x) &= \exp\{-x^{-\alpha}\}, \quad x > 0, \quad \alpha > 0.\end{aligned}$$

Redefining the constants  $a_n > 0$  and  $b_n \in \mathbb{R}$ , these can in turn be nested in a one-parameter family of distributions, the Generalized Extreme Value (GEV) distribution with distribution function

$$G_\gamma(x) := \exp\{-(1 + \gamma x)^{-1/\gamma}\}, \quad 1 + \gamma x > 0, \quad \gamma \in \mathbb{R}.$$

We then say that  $F$  is in the (max-)domain of attraction of  $G_\gamma$  and use the notation  $F \in \mathcal{D}_M(G_\gamma)$ . For  $\gamma < 0$ ,  $\gamma = 0$  and  $\gamma > 0$ , the GEV distribution function reduces again to Weibull, Gumbel and Fréchet distribution functions, respectively. An equivalent extreme value condition allows the limit relation in (1) to run over the real line (cf. Theorem 1.1.6 de Haan and Ferreira, 2006):  $F \in \mathcal{D}_M(G_\gamma)$  if and only if

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + b(t))) = (1 + \gamma x)^{-1/\gamma}, \quad (2)$$

for all  $x$  such that  $1 + \gamma x > 0$ ,  $a(t) := a_{[t]}$  and  $b(t) := b_{[t]}$ , with  $[t]$  denoting the integer part of  $t$ . The extreme value index  $\gamma$  determines various degrees of tail heaviness. If  $F \in \mathcal{D}_M(G_\gamma)$  with  $\gamma > 0$ , then the distribution function  $F$  is heavy-tailed, i.e.,  $F$  has a power-law decaying tail with infinite right endpoint. On the opposite end,  $\gamma < 0$  refers to short tails which must have finite right endpoint. The Gumbel domain of attraction  $\mathcal{D}_M(G_0)$  encloses a great variety of distributions, ranging from light-tailed distributions such as the Normal distribution, the exponential distribution, to moderately heavy distributions such as the Lognormal. All the just mentioned distributions have an infinite right endpoint but a finite endpoint is also possible in the Gumbel domain. We shall give several examples in Section 2. Distribution functions of this sort, i.e. light-tailed distributions with finite endpoint, but not so light that they are still included in the Gumbel domain, have been in great demand as feasible distributions underlying real life phenomena. A striking example is the extreme value analysis by Einmahl and Magnus (2008) of the best marks in Athletics, aiming at assessing the ultimate records for several events. For instance, Table 3 in Einmahl and Magnus (2008) has several missing values for the estimates of the endpoint which are due to an estimated extreme value index  $\gamma$  near zero. An attempt to fulfill these blank spaces with an appropriate

framework for inference in the Gumbel domain has been provided by Fraga Alves et al. (2013), although from the strict view point of application to the Long Jump data set used in Einmahl and Magnus (2008). The tentative estimator proposed by Fraga Alves et al. (2013) is virtually the same as the one introduced in the present paper. The novelty here is in the development of a simple closed-form expression for the previous statistic. Hence, the problem of estimating the right endpoint  $x^F$  of a distribution function lying in the Gumbel extremal domain of attraction is now tackled by the semi-parametric statistic

$$X_{n,n} + X_{n-k,n} - \frac{1}{\log 2} \sum_{i=0}^{k-1} \log\left(\frac{k+i+1}{k+i}\right) X_{n-k-i,n},$$

or in a more compact form, by

$$\hat{x}^F := X_{n,n} + \sum_{i=0}^{k-1} a_{i,k} (X_{n-k,n} - X_{n-k-i,n}), \quad (3)$$

where  $a_{i,k} := -(\log 2)^{-1}(\log(k+i) - \log(k+i+1)) > 0$ , such that  $\sum_{i=0}^{k-1} a_{i,k} = 1$ . Here and throughout this paper, the number  $k$  is assumed intermediate, that is,  $k$  is in fact a sequence of positive integers going to infinity as  $n \rightarrow \infty$  but at a much slower rate than  $n$ . More formally, we are defining  $\hat{x}^F$  as a functional of the top observations of the original sample, relying on an intermediate sequence  $k = k_n$ , i.e.

$$k_n \rightarrow \infty, \quad k_n = o(n), \quad \text{as } n \rightarrow \infty.$$

From the non-negativeness of the weighted spacings in the sum (3), we clearly see that the now proposed estimator  $\hat{x}^F$  is greater than the maximum  $X_{n,n}$  with probability one. This constitutes a crucial advantage in comparison with the usual semi-parametric estimators for the right endpoint of a distribution function in the Weibull domain of attraction (i.e. with  $\gamma < 0$ ). We refer to Hall (1982), Falk (1995), Hall and Wang (1999) and to de Haan and Ferreira (2006) and references therein. To the best of our knowledge, none of these estimators have ensured so far the extrapolation beyond the sample range, meaning that we can encounter in practice estimates for the endpoint smaller than the observed sample maximum. There have been, however, some developments of the most well-known endpoint estimators connected with  $\gamma < 0$  in the sense of bias reduction and/or correction. Li and Peng (2009), Li et al. (2011) and Cai et al. (2012) are a few of the most recent works in this respect. In fact, the problem of estimating  $x^F$  still gathers a great in-

terest nowadays. Recently, Girard et al. (2012) devised an endpoint estimator from the high-order moments pertaining to a distribution attached with  $\gamma < 0$ ; Li and Peng (2012) proposed a bootstrap estimator for the endpoint evolving from the one by Hall (1982) in case  $\gamma \in (-1/2, 0)$ . The present paper deliberately addresses the class of distribution functions belonging to the Gumbel domain of attraction, for which no specific inference has yet been provided in the context of estimation of the right endpoint  $x^F < \infty$ . The appropriate framework for the latter shall be developed in Section 2.

The remainder of the paper is as follows. The rationale behind the proposal of the new estimator for the right endpoint is expounded in Section 3. Large sample properties of this estimator, namely consistency and asymptotic distribution, are worked out in Section 4 by taking advantage of this form of separability between the maximum and the sum of higher order statistics. In order to perform asymptotics, we require some basic conditions in the context of the theory of regular variation. These are laid out in the next section (Section 2). In Section 5 we gather some simulation results taken as key examples. Finally, Section 6 is devoted to some applications and drawing out conclusions.

## 2 Framework

Let  $F$  be a distribution function (d.f.) with right endpoint  $x^F$ ,

$$x^F := \sup\{x : F(x) < 1\}.$$

For now we are assuming  $x^F \leq \infty$ .

Suppose  $F \in \mathcal{D}_M(G_\gamma)$ , that is  $F$  satisfies the following extreme value condition

$$\lim_{x \uparrow x^F} \frac{1 - F(t + x f(t))}{1 - F(t)} = (1 + \gamma x)^{-1/\gamma}, \quad (4)$$

for all  $x \in \mathbb{R}$  such that  $1 + \gamma x > 0$ , with a suitable positive function  $f$  (equivalent condition to (2), see Theorem 1.1.6 of de Haan and Ferreira, 2006).

For the most interesting case of  $\gamma = 0$  the limit in (4) reads as  $e^{-x}$ . In this case  $f > 0$  can be defined as follows

$$f(t) := \int_t^{x^F} \frac{1 - F(x)}{1 - F(t)} dx = E[X - t | X > t] \quad (5)$$

(cf. Theorem 1.2.5 of de Haan and Ferreira, 2006), then  $f$  is the so called Mean Excess Function.

Now let  $U$  be the (generalized) inverse function of  $1/(1 - F)$ . If  $F$  satisfies (4) with  $\gamma = 0$  then

we can assume there exists a positive function  $a_0$  such that, for all  $x > 0$ ,

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a_0(t)} = \log x. \quad (6)$$

Hence  $U$  belongs to the class  $\Pi$  (see Definition B.2.4 of de Haan and Ferreira, 2006) and  $a_0$  is a measurable function such that  $\lim_{t \rightarrow \infty} a_0(tx)/a_0(t) = 1$  for all  $x > 0$ . Then we say that  $a_0$  is a slowly varying function and use the notation  $a_0 \in RV_0$  (see Theorem B.2.7 of de Haan and Ferreira, 2006). Moreover, the functions  $a_0$  and  $f$  (introduced in (6) and (4), respectively) are related to each other by  $a_0 = f \circ U$  (see Theorem B.2.21 of de Haan and Ferreira, 2006). Here and throughout the paper, we use the notation  $U \in \Pi(a_0)$  in order to put some emphasis on the auxiliary function  $a_0$ . We shall assume the following conditions:

(A1)  $U \in \Pi(a_0)$ ;

(A2)  $U(t) = U(t_0) + \int_{t_0}^t a(s) \frac{ds}{s} + o(a(t))$ , for some  $t_0 \geq 1$ , with a positive function  $a \in RV_0$  satisfying  $a(t) \sim a_0(t)$ , as  $t \rightarrow \infty$ .

(B)  $x^F := U(\infty) = \lim_{t \rightarrow \infty} U(t)$  exists finite.

Assuming (A1), then Proposition B.2.15(3) of de Haan and Ferreira (2006) guarantees the existence of a twice differentiable function  $\bar{f}$ , with  $-\bar{f}'' \in RV_{-2}$ , such that  $U(t) = \bar{f}(t) + o(a_0(t))$ . Let  $\bar{f}(t) = \bar{f}(t_0) + \int_{t_0}^t \bar{f}'(s) ds$  be this function. Hence,  $U(t) = U(t_0) + \bar{f}(t) - \bar{f}(t_0) + o(a(t))$ , with  $a(t) \sim a_0(t)$  and where we set  $\bar{f}'(t) = a(t)/t$ . This is condition A2. Conversely, (A2) implies (A1) by Proposition B.2.15(5) of de Haan and Ferreira (2006) with  $g(s) = a(s)/s \in RV_{-1}$  therein.

(C) The present development (i.e. assuming (A2) and (B)) gives rise to

$$U(\infty) - U(t) = \int_t^\infty a(s) \frac{ds}{s} + o(a(t)), \quad t \rightarrow \infty, \quad (7)$$

which is our main assumption eventually. We note that by writing  $U(\infty)$  we are automatically assuming  $x^F = U(\infty)$  finite (i.e., condition B is embedded in condition C). Defining

$$q(t) := \int_t^\infty a(s) \frac{ds}{s} = \int_1^\infty a(st) \frac{ds}{s} = \int_0^1 a\left(\frac{t}{s}\right) \frac{ds}{s}, \quad (8)$$

then (C) rephrases as  $U(\infty) - U(t) = q(t) + o(a(t))$ , as  $t \rightarrow \infty$ .

We can obtain from (6) with  $a_0$  replaced by  $a$  (i.e.  $U \in \Pi(a)$ ) yet another limiting relation now

involving integration of  $U$  and  $a$ : applying Cauchy's rule once, we obtain

$$\lim_{t \rightarrow \infty} \frac{\int_0^{1/t} \left( U\left(\frac{x}{s}\right) - U\left(\frac{1}{s}\right) \right) \frac{ds}{s}}{\int_0^{1/t} a\left(\frac{1}{s}\right) \frac{ds}{s}} = \lim_{t \rightarrow \infty} \frac{(U(tx) - U(t))/t}{a(t)/t}, \quad (9)$$

then for arbitrary positive  $x$ , the  $\Pi$ -variation of  $U$  ascertains that  $\log x$  is the limit above, i.e.

$$\lim_{t \rightarrow \infty} \frac{\int_{tx}^{\infty} U(s) \frac{ds}{s} - \int_t^{\infty} U(s) \frac{ds}{s}}{\int_t^{\infty} a(s) \frac{ds}{s}} = \log x, \quad (10)$$

for all  $x > 0$ . Hence  $\int_t^{\infty} U(s) ds/s$  is also  $\Pi$ -varying with auxiliary function  $q$  introduced in (8). In the usual notation, the latter is  $\int_t^{\infty} U(s) ds/s \in \Pi(q)$ . Then  $q$  is slowly varying while condition C entails that  $q(t) \rightarrow 0$  as  $t \rightarrow \infty$  (cf. Lemma 15 from Appendix C).

Some examples of distributions belonging to the Gumbel domain of attraction with finite right endpoint, i.e., such that the main condition (7) holds, are listed below.

**Example 1** A random variable  $X$  is Negative Fréchet with parameter  $\beta > 0$  if it has distribution function  $F(x) = 1 - \exp\{-(x^F - x)^{-\beta}\}$ ,  $x \leq x^F$ ,  $\beta > 0$ . The associated tail quantile function  $U$  is given by  $U(t) = F^{\leftarrow}(1 - 1/t) = x^F - (\log t)^{-1/\beta}$ ,  $t \geq 1$  (the arrow stands for the generalized inverse). Then  $U \in \Pi(a_0)$  with  $a_0(t) = (1/\beta)(\log t)^{-1/\beta-1} \rightarrow 0$ , as  $t \rightarrow \infty$ . Therefore, the auxiliary function  $q$  in (10) becomes  $q(t) = (\log t)^{-1/\beta}$ ,  $\beta > 0$ .

**Example 2** Consider the distribution function  $F$  given by  $F(x) = 1 - \exp\{-\tan(x/\beta)\}$ ,  $0 \leq x < \beta\pi/2$ ,  $\beta > 0$ . The pertaining function  $U(t) = F^{\leftarrow}(1 - 1/t)$  is given by  $U(t) = \beta \arctan(\log t)$ ,  $t \geq 1$ . Hence  $U$  satisfies the main condition (7) (condition C) with  $a(t) = 1/(\log^2 t + \beta^{-2})$  and  $U \in \Pi(a)$  where  $U(\infty) = \beta\pi/2 = x^F$ .

**Example 3** Consider the distribution function  $F(x) = 1 - \exp\{(\pi/2)^{-\beta} - (\arcsin(1 - x/\beta))^{-\beta}\}$ ,  $0 \leq x < \beta$ ,  $\beta > 0$ . Then  $U(t) = \beta\{1 - \sin([(2/\pi)^\beta + \log t]^{-1/\beta})\}$ ,  $t \geq 1$ . Therefore, condition C holds with  $a(t) = (\log t)^{-(1/\beta+1)} \cos((\log t)^{-1/\beta})$ ,  $U \in \Pi(a)$  and  $U(\infty) = \beta = x^F$ .

### 3 Statistics

Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from the underlying distribution function  $F$  with finite right endpoint  $x^F$ . Let  $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$  be the corresponding order statistics. We introduce the estimator  $\hat{q}(n/k)$  for the auxiliary function  $q$  defined in (8), evaluated at  $t = n/k$ .

This estimator has the property that, as  $n \rightarrow \infty$ ,  $k = k(n) \rightarrow \infty$  and  $k(n)/n \rightarrow 0$  (provided some suitable yet mild restrictions involving the second order refinement of  $\int_t^\infty U(s)/s ds$ ),

$$\frac{q\left(\frac{n}{k}\right)}{a\left(\frac{n}{k}\right)} \left( \frac{\hat{q}\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} - 1 \right) \xrightarrow[n \rightarrow \infty]{d} N,$$

where  $N$  is a non-degenerate random variable. Several estimators for the right endpoint  $x^F = U(\infty) < \infty$  can be readily devised from (7), in the sense that these might evolve from a suitable estimator  $\hat{q}(n/k)$  for  $q(n/k)$ , i.e.

$$\hat{x}^F = \hat{U}\left(\frac{n}{k}\right) + \hat{q}\left(\frac{n}{k}\right) = X_{n-k,n} + \hat{q}\left(\frac{n}{k}\right). \tag{11}$$

The above relation allows to foresee that  $\hat{x}^F$  carries analogous large sample properties to  $\hat{q}(n/k)$ . In particular, the consistency of  $\hat{x}^F$  is essentially ensured by the consistency of  $\hat{q}(n/k)$ . Theorem 5 in Section 4 accounts for this.

We now evaluate relation (10) at  $x = 1/2$  together with  $q(t)$  at  $t = n/k$  (see last equality in (8)). This prompts the following approximation for large enough  $n$ :

$$\int_0^1 \left( U\left(\frac{n}{2ks}\right) - U\left(\frac{n}{ks}\right) \right) \frac{ds}{s} \approx q\left(\frac{n}{k}\right) (-\log 2).$$

Our proposal for estimating  $q(n/k)$  thus arises quite naturally from the empirical counterparts  $\hat{U}(n/(\theta ks)) = X_{n-[\theta ks],n}$ ,  $s \in (0, 1]$ ,  $\theta = 1, 2$ , i.e.

$$\hat{q}\left(\frac{n}{k}\right) := -\frac{1}{\log 2} \int_0^1 (X_{n-[2ks],n} - X_{n-[ks],n}) \frac{ds}{s}. \tag{12}$$

A certain amount of simple calculations yields the following alternative expression for  $\hat{q}$ :

$$\hat{q}\left(\frac{n}{k}\right) = X_{n,n} + \frac{1}{\log 2} \sum_{i=0}^{k-1} \log\left(\frac{k+i}{k+i+1}\right) X_{n-k-i,n}. \tag{13}$$

Combining (11) with (13) we are led to the estimator for the right endpoint

$$\hat{x}^F := X_{n-k,n} + X_{n,n} + \frac{1}{\log 2} \sum_{i=0}^{k-1} \log\left(\frac{k+i}{k+i+1}\right) X_{n-k-i,n}. \tag{14}$$

We note that, after rearranging components, it is possible to express  $\hat{x}^F$  as the maximum  $X_{n,n}$



added by some weighted mean of non-negative summands as follows:

$$\hat{x}^F = X_{n,n} + \sum_{i=0}^{k-1} a_{i,k} (X_{n-k,n} - X_{n-k-i,n}),$$

with  $a_{i,k} := -(\log 2)^{-1}(\log(k+i) - \log(k+i+1)) > 0$ ,  $i = 1, 2, \dots, k \in \mathbb{N}$ , such that  $\sum_{i=0}^{k-1} a_{i,k} = 1$ .

**Remark 4** We emphasize that the now proposed estimator for the right endpoint returns values always larger than  $x_{n,n}$ . This constitutes a major advantage in comparison to the available semi-parametric estimators for the endpoint in the case of Weibull domain of attraction, for which the extrapolation beyond the sample range is not guaranteed. This inadequacy of the existing estimators often leads to some disappointing results in practical applications, with estimates-yields that may be lower than the observed maximum in the data.

## 4 Asymptotic results

In this section we shall bear in mind that  $\hat{x}^F$  is a compound of the two building blocks introduced in (11): the sufficiently high random threshold  $X_{n-k,n}$  and  $\hat{q}(n/k)$  defined in (12). Our line of reasoning pursues the fact that the large sample properties of the new estimator  $\hat{x}^F$  are essentially governed by the asymptotic properties of the estimator  $\hat{q}(n/k)$ . This is the statement in our main Theorem 5 with respect to consistency. A proof is provided afterwards in this Section. Hence, the consistency of our estimator for the right endpoint stems from the consistency of  $\hat{q}(n/k)$ . The latter is tackled in Appendix A, with  $\hat{q}(n/k)$  defined in (12) (see also (13)), for an intermediate sequence  $k = k_n$ . Similarly, the limiting distribution of  $\hat{q}(n/k)$  in Theorem 8 renders the asymptotic distribution of  $\hat{x}^F$  via Proposition 11. All the remainder proofs regarding  $\hat{q}(n/k)$  are postponed to Appendix B.

**Theorem 5** Let  $X_1, X_2, \dots$  be i.i.d. random variables with tail quantile function  $U$  satisfying condition C. Suppose  $k = k_n$  is a sequence of positive integers such that  $k_n \rightarrow \infty$ ,  $k_n/n \rightarrow 0$ , as  $n \rightarrow \infty$ , and  $\hat{q}(n/k)$  is a consistent estimator for  $q(n/k)$  in the sense that the following convergence in probability holds:  $\hat{q}(n/k)/q(n/k) \xrightarrow{P} 1$ .

Then  $\hat{x}^F := X_{n-k,n} + \hat{q}(n/k)$  is a consistent estimator for  $x^F < \infty$ , i.e.  $\hat{x}^F \xrightarrow[n \rightarrow \infty]{P} x^F$ .

**Proof:** Let  $U_{k+1,n}$  be the  $(k+1)$ -th ascending order statistic from the random sample  $(U_1, U_2, \dots, U_n)$  of  $n$  uniformly distributed random variables on the unit interval. Then  $X_{n-k,n} \stackrel{d}{=} U(1/U_{k+1,n})$ , where  $U(\cdot)$  denotes the underlying tail quantile function and  $\stackrel{d}{=}$  stands for equality in distribution,

as usual. It will suffice to note there are three main contributing components for  $x^F - \hat{x}^F$ , but these become negligible with increasing  $n$ . Specifically,

$$\begin{aligned} x^F - \hat{x}^F &\stackrel{d}{=} \left( U(\infty) - U\left(\frac{n}{k}\right) - q\left(\frac{n}{k}\right) \right) - \left( U\left(\frac{1}{U_{k+1,n}}\right) - U\left(\frac{n}{k}\right) \right) - q\left(\frac{n}{k}\right) \left( \frac{\hat{q}\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} - 1 \right) \\ &= I - II - III, \end{aligned}$$

where:

$$I := U(\infty) - U\left(\frac{n}{k}\right) - q\left(\frac{n}{k}\right) = o\left(a\left(\frac{n}{k}\right)\right),$$

which follows directly from relation (7);

$$II := U\left(\frac{1}{U_{k+1,n}}\right) - U\left(\frac{n}{k}\right) = o_p\left(a\left(\frac{n}{k}\right)\right)$$

because  $U \in \Pi(a)$  while Smirnov's Lemma ensures  $k/(nU_{k+1,n}) \xrightarrow{P} 1$  (see Lemma 2.2.3 in de Haan and Ferreira, 2006);

$$III := q\left(\frac{n}{k}\right) \left( \frac{\hat{q}\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} - 1 \right) = o_p(1)$$

which is verified by Proposition 13 and the fact that relation (7) implies  $q(n/k) = o(1)$ .  $\square$

The limiting distribution of  $\hat{q}(n/k)$  (and, later on, the asymptotic distribution of  $\hat{x}^F$ ) is attained under a suitable second order refinement of (6): suppose there exist functions  $a$ , positive and  $A$ , positive or negative, both tending to zero as  $t \rightarrow \infty$ , such that

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx) - U(t)}{a(t)} - \log x}{A(t)} = \frac{1}{2}(\log x)^2, \quad (15)$$

for all  $x > 0$ .

**Remark 6** *The second order condition above follows directly from Theorem B.3.6, Remark B.3.7 and Corollary 2.3.5 of de Haan and Ferreira (2006) because the former states that, in our setup of  $\gamma = 0$  and  $x^F < \infty$ , the only case allowed is the case of the second order parameter  $\rho$  equal to zero. Like the function  $a$ , the second order auxiliary function  $A$  converges to zero, not changing sign for  $t$  near infinity, and  $|A|$  is slowly varying, i.e.  $A(tx)/A(t) \rightarrow 1, t \rightarrow \infty$  (notation:  $|A| \in RV_0$ ).*

**Example 7** *We consider again the Negative Fréchet model with parameter  $\beta > 0$ , i.e., , with distribution function  $F(x) = 1 - \exp\{-(x^F - x)^{-\beta}\}$ ,  $x \leq x^F$ ,  $\beta > 0$  (see Example 1). The pertaining auxiliary function is  $q(t) = (\log t)^{-1/\beta}$ ,  $\beta > 0$ . Now, straightforward calculations yield*

$A_0(t) = -(1 + 1/\beta)(\log t)^{-1}$ , which implies that  $-a_0(t)/q(t) = A_0(t)/(1 + \beta)$ , for  $t$  near infinity.

As announced in the beginning of this section, given the two main building blocks of our estimator for the right endpoint, the next theorem enables the asymptotic distribution of  $\hat{x}^F$  stated in Proposition 11. The proof of Theorem 8 can be found in Appendix B.

**Theorem 8** *Given condition C (i.e. suppose relation (7) holds), assume the second order condition (15) holds. Suppose  $k = k_n$  is such that, as  $n \rightarrow \infty$ ,  $k_n \rightarrow \infty$ ,  $k_n/n \rightarrow 0$ ,  $a(n)/a(n/k_n) \rightarrow 1$  and  $\sqrt{k_n} A(n/k_n) = O(1)$ . Furthermore assume that*

$$\lim_{n \rightarrow \infty} \frac{1}{A(n/k)} \left( \int_{\frac{1}{2k}}^1 \frac{U(\frac{n}{ks}) - U(\frac{n}{2ks})}{q(\frac{n}{k})} \frac{ds}{s} - \log 2 \right) = \lambda \in \mathbb{R}. \quad (16)$$

Then

$$\frac{q(\frac{n}{k})}{a(\frac{n}{k})} \left( \frac{\hat{q}(\frac{n}{k})}{q(\frac{n}{k})} - 1 \right) \xrightarrow[n \rightarrow \infty]{d} \Lambda - \frac{\log 2}{2} - \frac{\lambda}{\log 2}, \quad (17)$$

where  $\Lambda$  is a Gumbel random variable with distribution function  $\exp\{-e^{-x}\}$ , for all  $x \in \mathbb{R}$ .

We note that the assumption (16) of the theorem regards a second order refinement of (10), more concretely:

$$\lim_{t \rightarrow \infty} \frac{\int_{tx}^{\infty} U(s) \frac{ds}{s} - \int_t^{\infty} U(s) \frac{ds}{s} - \log x}{Q(t)} = \frac{1}{2}(\log x)^2, \quad (18)$$

taken in the point  $x = 2$  for large enough  $t = n/k$ . Hence, the assumption (16) has been tailored with  $Q(t) = O(A(t))$ , via the usual second order setup (see also Eq. (15)) provided by the theory of extended regular variation. We refer to Appendix B of de Haan and Ferreira (2006) for a good catalog on results concerning theory of extended regular variation.

The assumption on that  $a(n/k)/a(n) \rightarrow 1$ , as  $n \rightarrow \infty$ , is however, a bit more restrictive in terms of screening for an adequate value  $k$  which will determine the number of top order statistics to base our inference from. For example, if we assume the Negative Fréchet for the underlying distribution function (see Example 1) and  $k_n = n^p$ ,  $p \in (0, 1)$ , then

$$\frac{a(n)}{a(n/k_n)} = \left( 1 - \frac{\log k_n}{\log n} \right)^{1/\beta+1} = (1 - p)^{1/\beta+1},$$

which is approximately 1 if and only if  $p$  approaches zero. A more appropriate choice regards intermediate sequences at a slower rate such as  $k_n = (\log n)^r$ ,  $r \in (0, 2]$ . Bearing this choice in

mind, we have that

$$\frac{a(n)}{a(n/k_n)} = \left(1 - \frac{\log k_n}{\log n}\right)^{1/\beta+1} = \left(1 - \frac{r}{\log n} + \frac{\log \log n}{\log n}\right)^{1/\beta+1} \xrightarrow{n \rightarrow \infty} 1.$$

The upper bound  $r \leq 2$  is imposed in order to comply with the assumption  $\sqrt{k_n}A(n/k_n) = O(1)$ .

Given the slow variation feature of all the functions involved in the characterization of the present subclass of distributions in the Gumbel domain with finite right endpoint, we believe that the latter choice for  $k = k_n$  is a feasible one for most models satisfying (7), meaning that we require intermediate values  $k_n$  such that  $\log(k_n) = o(\log n)$ . Nevertheless, we can bring forward the fact that a miss-specification of  $k_n$  in the sense that  $a(n/k_n)/a(n)$  converges to a constant different than 1, has a direct impact on the asymptotic variance of the normalized relative error presented in Theorem 8 rather than upon the asymptotic bias. This can be clearly seen in the proof of Theorem 8. In this respect, we defer the reader to Appendix B.

**Remark 9** We note that the assumption  $\log k = o(\log n)$  is a common assumption in the theoretical analysis of estimators for Weibull-type tails, which form a rich subclass the Gumbel max-domain of attraction, albeit with  $x^F = \infty$ . In this respect, we refer to Goegebeur et al. (2010) and Gardes et al. (2011).

**Example 10** We resume the results for the Negative Fréchet distribution introduced in Example 1 (see Example 7 on further developments). The Negative Fréchet distribution has a tail quantile function given by  $U(t) = x^F - (\log t)^{-1/\beta}$ ,  $t \geq 1$ ,  $0 < \beta < 1$ , satisfying the second order condition (18) with  $Q(t) = -(\beta \log t)^{-1}$ .

We are thus ready to proceed with the asymptotic distribution of  $\hat{x}^F$ . The following proposition rests heavily on the statement in Theorem 8.

**Proposition 11** Under the conditions of Theorem 8,

$$\frac{1}{a(n/k)} (\hat{x}^F - x^F) - \frac{q(\frac{n}{k})}{a(\frac{n}{k})} \left( \frac{\hat{q}(\frac{n}{k})}{q(\frac{n}{k})} - 1 \right) \xrightarrow{P} 0.$$

**Proof:** We write:

$$\begin{aligned} \frac{\hat{x}^F - x^F}{a(n/k)} - \frac{q(n/k)}{a(n/k)} \left( \frac{\hat{q}(n/k)}{q(n/k)} - 1 \right) &= \frac{\hat{x}^F - \hat{q}(n/k)}{a(n/k)} - \frac{x^F - q(n/k)}{a(n/k)} \\ &= \frac{X_{n-k,n} - U(n/k)}{a(n/k)} - \frac{U(\infty) - U(n/k) - q(n/k)}{a(n/k)}. \end{aligned}$$

Since under the second order condition (15), Theorem 2.4.1 of de Haan and Ferreira (2006) ensures that  $(X_{n-k,n} - U(n/k))/a(n/k) = O_p(1/\sqrt{k}) = o_p(1)$ , the rest follows from condition C.  $\square$

Finally, Theorem 12 below encloses an alternative formulation of the results comprised in Theorem 8 and Proposition 11 aiming at providing confidence bands for  $\hat{x}^F$ .

**Theorem 12** *Let  $X_1, X_2, \dots$  be i.i.d. random variables with tail quantile function  $U$  satisfying the second order condition (15). Let  $\hat{a}(n/k)$  be a consistent estimator for  $a(n/k)$ . Suppose  $k = k_n$  is a sequence of positive integers such that, as  $n \rightarrow \infty$ ,  $k_n \rightarrow \infty$ ,  $k_n/n \rightarrow 0$ ,  $a(n)/a(n/k_n) \rightarrow 1$  and  $\sqrt{k_n} A(n/k_n) = O(1)$ . Furthermore assume that*

$$\lim_{n \rightarrow \infty} \frac{1}{A(n/k)} \left( \int_{\frac{1}{2k}}^1 \frac{U(\frac{n}{ks}) - U(\frac{n}{2ks})}{q(\frac{n}{k})} \frac{ds}{s} - \log 2 \right) = \lambda \in \mathbb{R}.$$

Then

$$\frac{1}{\hat{a}(n/k)} (\hat{x}^F - x^F) \xrightarrow[n \rightarrow \infty]{d} \Lambda - \frac{\log 2}{2} - \frac{\lambda}{\log 2}.$$

**Proof:** The result follows easily by conjugating Theorem 8 with Proposition 11 and then applying Slutsky's theorem.  $\square$

There are in the literature several possibilities for estimating the auxiliary (or scale) function  $a(n/k)$ . The most obvious choice is the Maximum Likelihood Estimator (MLE) by pretending that the exceedances over a certain high (random) threshold follow a Generalized Pareto distribution (cf. section 3.4 of de Haan and Ferreira, 2006):

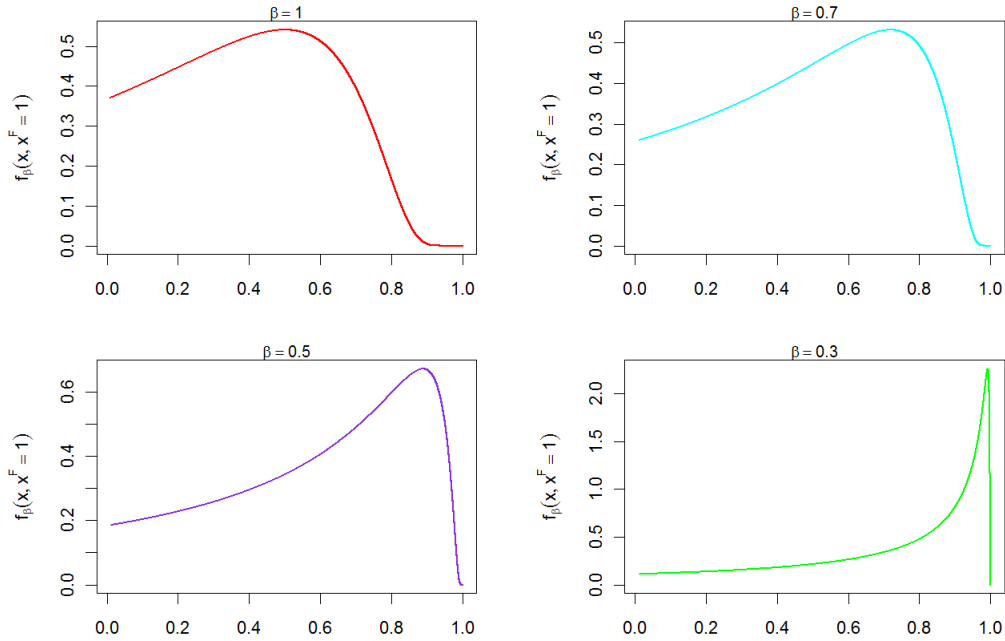
$$\hat{a}\left(\frac{n}{k}\right) = \hat{\sigma}^{MLE} := \frac{1}{k} \sum_{i=0}^{k-1} (X_{n-i,n} - X_{n-k,n}).$$

## 5 Simulations

The Negative Fréchet distribution features in this simulation for the purpose of illustrating the finite sample behavior of our estimator for  $x^F$  defined in (14). Other models would have been equally possible, for instance the ones provided in Examples 2 and 3.

Negative Fréchet model, with distribution function  $F(x) = 1 - \exp\{-(x^F - x)^{-\beta}\}$ ,  $x \leq x^F$ ,  $\beta > 0$ , reveals a tail shape versatility, for different values of  $\beta$ . The pertaining tail quantile function  $U$  is given by  $U(t) = x^F - (\log t)^{-1/\beta}$ ,  $t \geq 1$ . Clearly  $U \in \Pi(a)$  with  $a(t) = \beta^{-1}(\log t)^{-1/\beta-1}$ ,  $\beta > 0$

**Figure 1:** Right tails for probability density functions of Negative Fréchet Model with right endpoint  $x^F = 1$  and  $\beta = 0.3, 0.5, 0.7, 1$ .

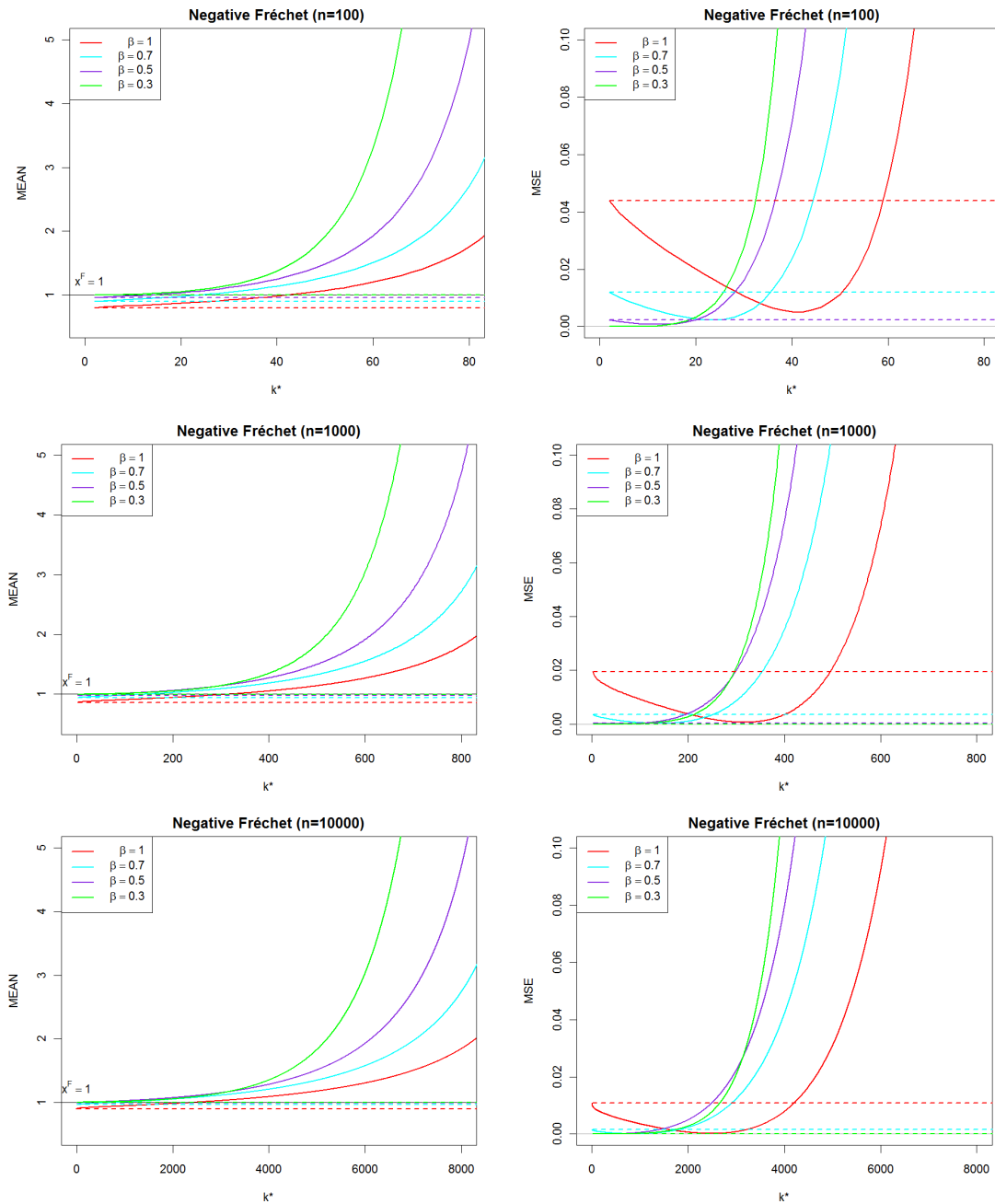


(cf. Example 1). The range of the chosen values for  $\beta$  offers various tail shapes, as shown in the graphics drawn in Figure 1.

We have simulated 1000 samples of size  $n = 100, 1000, 10000$ , from Negative Fréchet model with right endpoint  $x^F = 1$  and for different parameters  $\beta = 0.3, 0.5, 0.7, 1$ . The results are depicted in Figure 2. Since the number  $k$  actually implies that the number of top order statistics used in the estimation is twice as much, we have plotted the estimated mean of  $\hat{x}^F$  as a function of the latter, i.e., the plots are against  $k^* = 2k$ , from 1 to 80%  $n$ .

The most common approach of selecting the number  $k$  (or  $k^*$  in the present case) is to look for a region where the plots are relatively stable. This way, given the consistency property of the adopted estimator, one should in principle be away from small values of  $k$  avoiding large variance (small  $k$  is usually associated with a large variance) and not so far off in the tail preventing bias to instill (bias usually due to large  $k$ ). As already discussed in Section 4, for Model 1 an appropriate choice for an intermediate  $k = k_n$  may be given by  $k_n = (\log n)^r$ , with  $r \in (0, 2]$ . If we are using  $n = 1000$ , for instance, and if we set  $r = 2$ , the maximum allowed for  $r$ , we obtain  $k \approx 48$  and thus  $k^* \approx 96$ . Bearing on a value of  $k^*$ , around 100 e.g., all the plots in Figure 2 look quite stable in a

Figure 2: Mean estimate (left) and empirical Mean Squared Error (right) of  $\hat{x}^F$  defined in (14) (solid line), for Negative Fréchet Model with the true value  $x^F = 1$ ,  $\beta = 0.3, 0.5, 0.7, 1$  and several sample sizes:  $n = 100$  (first row),  $n = 1000$  (second row),  $n = 10000$  (third row); All plots are depicted against the number  $k^* = 2k$  of top observations used in the estimator. The naive maximum estimator,  $\tilde{x}^F := X_{n,n}$ , has also been considered (dashed line).



close vicinity of the target value  $x^F = 1$  represented by the solid horizontal black line. Hence, the slow convergence imposed by intermediate sequence  $k_n = (\log n)^2$  seems to be of little effect to almost none, upon the finite sample behavior of the new estimator  $\hat{x}^F$ . The latter is particular true in case  $0 < \beta < 1$ .

A more thorough examination of the graphs in Figure 2 seems to give accounts of a tendency to a better estimation under Negative Fréchet model if the parameter  $\beta$  is less than 1, which corresponds to the case where the inherent second order conditions are satisfied. We recall that if  $\beta \geq 1$ , the Negative Fréchet distribution still satisfies the first order condition. Further details on the Negative Fréchet distribution are given in Examples 1, 7 and 10. Moreover, for this model, the general pattern for the mean estimate of  $\hat{x}^F$  involves a moderated bias with  $k^*$  in the upper part of the sample, and a fast increasing bias with  $k^*$  around 40% of the sample size.

Similar simulations have been carried out for the other two models in Examples 2 and 3, leading also to favorable results. Note that for any model with right endpoint finite, the sample path of  $\hat{x}^F$  departures from the top value  $x_{n,n}$ , i.e., , the sample maximum, and always returns extrapolated values beyond the sample. We should highlight that our estimator  $\hat{x}^F$  yields better results than the maximum estimator  $\tilde{x}^F := X_{n,n}$ , which always underestimates the true value  $x^F$ . The relative performance of both  $\tilde{x}^F$  and  $\hat{x}^F$  can also be easily observed if we compare the MSE graphics in Figure 2 (*right*): for the top part the sample, depending on the  $\beta$  value, the estimator  $\hat{x}^F$  always outperforms the maximum  $\tilde{x}^F$ , presenting the new estimator a lower mean squared error than the naive maximum estimator.

## 6 Case study and conclusion

This section is dedicated to the estimation of the finite right endpoint in the Gumbel maximum domain of attraction, which embraces light-tailed distributions with finite endpoint. This topic has been developed under the motivation of real life data problems, where extreme value analysis demands the estimation of the right endpoint, although the underlying distribution tail is not so light that could be included in Weibull domain of attraction, i.e., where the assumption of negative extreme value index (EVI),  $\gamma < 0$ , is questionable.

A seminal approach has been worked out in Fraga Alves et al. (2013), with a real example applied on athletics records data, filling the gap highlighted in Einmahl and Magnus (2008), on extreme value analysis of the best marks in Athletics, aiming at assessing the ultimate records for several events. In particular, in Einmahl and Magnus (2008) and for the Long Jump event (LJ-



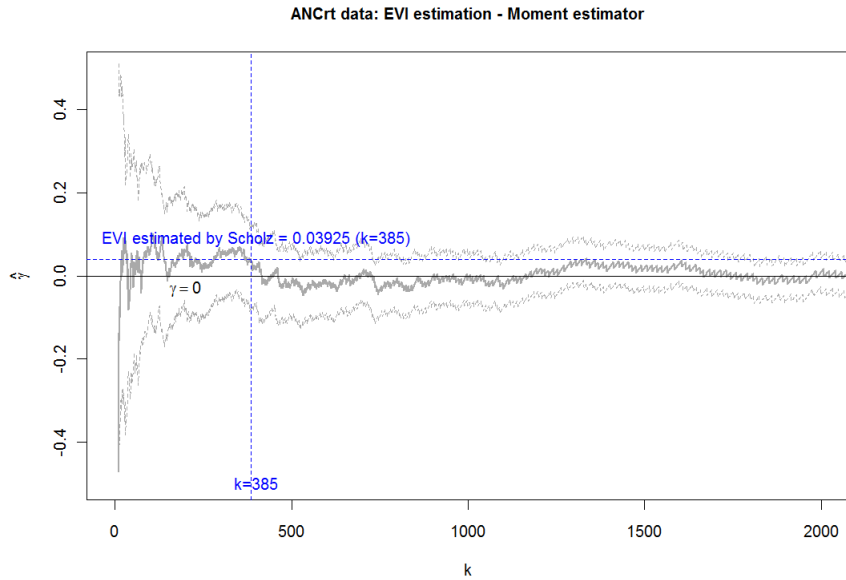
Men), the authors estimated a positive EVI much close to zero, which compromises the endpoint estimation with the available methodology at the time. In Fraga Alves et al. (2013), and for LJ data taken in the same period of time, an answer to the applied problem of estimation the right endpoint of LJ-Men event has been given, using this same estimator  $\hat{x}^F$ , but written in an equivalent algebraic form; as referred therein, that contribution constitutes a first approach to the estimation of the right endpoint for a distribution in the Gumbel max-domain of attraction, at that time a subject not yet addressed in the literature of extremes. Although the estimator motivation is presented in Fraga Alves et al. (2013) by theoretical framework for inference in the Gumbel domain, encompassing specific tail properties for finite endpoint, the asymptotic properties of the endpoint estimator were still under study, and are finally addressed in the present paper, which can be seen as an important development in the subject.

Real data phenomena constitutes a rich portfolio for similar applied problems. We shall consider an application to statistical extreme value analysis of Anchorage International Airport (ANC) Taxiway Centerline Deviations for Boeing 747 Aircraft, handled by Scholz (2003).

The goal was to provide a basis for understanding the extreme behavior of centerline deviations of Boeing-747. That report addressed the risk of an aircraft deviating at a fixed location along the taxiway beyond a certain threshold distance from the taxiway centerline. The B-747 taxiway deviation data were collected from 9/24/2000 to 9/27/2001 at ANC; during this period, 9767 deviations were recorded at ANC with a range of  $[-8.225, 8.863]$  feet, in both directions of the taxiways. Based on the extreme value limiting assumption, positive deviations (ANCrt data with sample size  $n = 4900$ ) were extrapolated using the  $k = 385$  most extreme deviations at ANC, the chosen  $k$  value of top observations to EVI estimation, namely  $\hat{\gamma} = 0.03925$  (cf. Scholz (2003) pages 43, 48, 64). In Figure 3 it is depicted the samplepath of  $\hat{\gamma}$ , the EVI estimate, using Moment estimator presented in Dekkers et al. (1989), along with the 95% confidence bandwidths, for the ANCrt data. It is easily checked that the straight line corresponding  $\gamma = 0$  is inside the confidence bandwidths for a very large upper part of the sample, (in the graphic  $k \leq 2000$ ). Consequently, the Gumbel domain of attraction cannot be discarded. Moreover, the testing procedures for detecting a finite right endpoint (cf. Neves and Pereira (2010)) suggest the presence of a distribution with finite right endpoint underlying the ANC deviation data.

In Figure 4 it is depicted the samplepath of our endpoint estimator  $\hat{x}^F$  against  $k^*$ . In the range of  $650 \leq k^* \leq 2300$  the graphic is quite stable. If we rely on that region we suggest for endpoint estimate a value approximately equal to 9.21 ft. We highlight that this estimation result is based

Figure 3: ANCRt data: EVI estimation with Moment estimator and 95% confidence bandwidths, plotted against  $k$



on the data set ANCRt in the period 9/24/2000 to 9/27/2001, and any conclusion can only be driven for the state of the art at that time; today’s knowledge of the ANC event would improve the endpoint estimation.

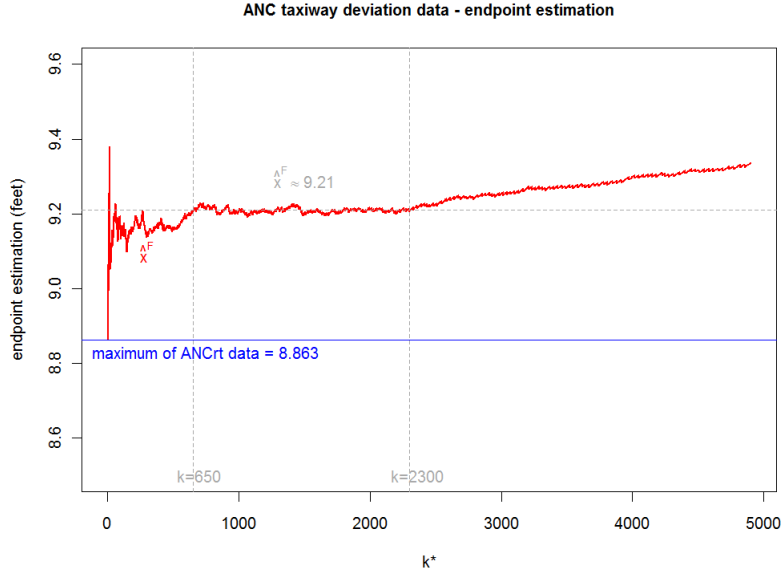
Taking all into account, we may conclude that the proposed estimator  $\hat{x}^F$  performs reasonably well for parent distributions in the Gumbel domain detaining finite right endpoint  $x^F$ .

As a short final remark about the robustness of endpoint estimator defined in (3), we can say it constitutes an advised inference procedure under Weibull domain of attraction. The theoretical background supporting this statement is a topic of further undergoing research, but beyond the scope of the present subject.

## Acknowledgement

The authors are grateful to Professor Laurens de Haan for introducing the appropriate characterization of distributions with finite right endpoint in the Gumbel domain, at the origin of the proposed estimator. We are grateful to Professor Fritz Scholz, who kindly shared the Anchorage Anchorage International Airport (ANC) data set.

Figure 4: endpoint estimation.



## A Consistency of $\hat{q}(n/k)$

Before we proceed we need to establish some ground results.

Let  $U_1, U_2, \dots, U_n$  be independent and identically distributed uniform random variables on the unit interval and let  $U_{1,n} \leq U_{2,n} \leq \dots \leq U_{n,n}$  be their order statistics. Since  $k = k_n$  is an intermediate sequence such that  $k_n \rightarrow \infty$ ,  $k_n = o(n)$ , as  $n \rightarrow \infty$ , then we can define a sequence of Brownian motions  $\{W_n(s)\}_{s \geq 0}$  such that, for each  $\varepsilon > 0$ ,

$$\sup_{\frac{1}{\theta k} \leq s \leq 1} s^{\frac{3}{2} + \varepsilon} \left| \sqrt{\theta k} \left( \frac{\theta k}{n U_{[\theta k s] + 1, n}} - \frac{1}{s} \right) - \frac{1}{s^2} W_n(s) \right| = o_p(1), \quad (19)$$

for all  $\theta \geq 1$  (cf. Lemma 2.4.10 of de Haan and Ferreira, 2006, with  $\gamma = 1$ ).

Let  $X_1, X_2, \dots$  be i.i.d random variables with the same distribution function  $F$  belonging to the Gumbel domain of attraction, i.e.,  $F \in \mathcal{D}(G_0)$ , with finite right endpoint  $x^F$ , such that condition C in Section 2 holds. Note that  $U(1/U_i) \stackrel{d}{=} X_i$ ,  $i = 1, 2, \dots$ . In view of condition C, the following relation holds

$$\frac{U(tx) - U(t)}{a(t)} = \int_{1/x}^1 \frac{a(\frac{t}{s})}{a(t)} \frac{ds}{s} + \frac{a(tx)}{a(t)} o(1) + o(1), \quad (t \rightarrow \infty)$$

for all  $x > 0$ . Given that  $a \in RV_0$ , we thus obtain for sufficiently large  $n$  that

$$\frac{X_{n-[ \theta ks ],n} - U\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta k}\right)} \stackrel{d}{=} \frac{U\left(\frac{n}{\theta k} \frac{\theta k}{nU_{[ \theta ks ]+1,n}}\right) - U\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta k}\right)} \approx \int_{\frac{nU_{[ \theta ks ]+1,n}}{\theta k}}^1 \frac{a\left(\frac{n}{\theta k} \frac{1}{x}\right)}{a\left(\frac{n}{\theta k}\right)} \frac{dx}{x}.$$

Now the uniform inequalities in Lemma 15.1(ii) tell us that, for any  $\varepsilon > 0$ ,

$$\frac{a\left(\frac{n}{\theta k} \frac{1}{s}\right)}{a\left(\frac{n}{\theta k}\right)} = 1 \pm \varepsilon s^{-\varepsilon}, 0 < s \leq 1.$$

Since  $U_{[ \theta ks ]+1,n} \in [0, 1]$  and for every  $s \in (0, 1]$ ,

$$\frac{nU_{[ \theta ks ]+1,n}}{\theta k} \leq \frac{nU_{[ \theta k ]+1,n}}{\theta k} \xrightarrow[n \rightarrow \infty]{P} 1,$$

we get the upper bound

$$\begin{aligned} \frac{X_{n-[ \theta ks ],n} - U\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta k}\right)} &\leq -\log s - \log\left(1 + \left(\frac{nU_{[ \theta ks ]+1,n}}{\theta ks} - 1\right)\right) + \left(\left(\frac{nU_{[ \theta ks ]+1,n}}{\theta k}\right)^{-\varepsilon} - 1\right) \\ &= -\log s - \frac{1}{s}\left(\frac{nU_{[ \theta ks ]+1,n}}{\theta k} - s\right)(1 + o_p(1)) + (s^{-\varepsilon} - 1)(1 + o_p(1)), \end{aligned}$$

with the  $o_p(1)$ -term tending to zero uniformly for  $s \in [(\theta k)^{-1}, 1]$ . A similar lower bound is also possible. Now we can apply Cramér's  $\delta$ -method to relation (19) in order to obtain:

$$\frac{X_{n-[ \theta ks ],n} - U\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta k}\right)} = -\log s + \frac{1}{\sqrt{\theta k}}\left(s^{-1}W_n(s) + o_p(s^{-1/2-\varepsilon})\right) \pm (s^{-\varepsilon} - 1)(1 + o_p(1)), \quad (20)$$

as  $n \rightarrow \infty$ , uniformly for  $(\theta k)^{-1} \leq s \leq 1$ ,  $\theta \geq 1$ . We now consider the normalized difference between a sample intermediate quantile and corresponding theoretical quantile and denote it by  $R_\theta(s)$ , i.e.

$$\begin{aligned} R_\theta(s) &:= \frac{X_{n-[ \theta ks ],n} - U\left(\frac{n}{\theta ks}\right)}{a\left(\frac{n}{\theta ks}\right)} \quad (21) \\ &= \frac{X_{n-[ \theta ks ],n} - U\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta k}\right)} + \left(\frac{a\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta ks}\right)} - 1\right) \frac{X_{n-[ \theta ks ],n} - U\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta k}\right)} + \frac{U\left(\frac{n}{\theta k}\right) - U\left(\frac{n}{\theta ks}\right)}{a\left(\frac{n}{\theta ks}\right)} \end{aligned}$$

Bearing on (20) combined with the uniform inequalities in Lemma 15(1), we thus get for any  $\varepsilon > 0$ ,

$$\begin{aligned} R_\theta(s) &= -\log s + \frac{1}{\sqrt{\theta k}} \left( \frac{W_n(s)}{s} + s^{-1/2-\varepsilon} o_p(1) \right) \\ &\quad \pm (s^{-\varepsilon} - 1)(1 + o_p(1)) \pm \varepsilon s^{-\varepsilon} (-\log s) + \log s \pm \varepsilon s^{-\varepsilon} \\ &= \frac{1}{\sqrt{\theta k}} \frac{W_n(s)}{s} \pm (s^{-\varepsilon} - 1)(1 + o_p(1)) \mp \varepsilon s^{-\varepsilon} \log s, \end{aligned} \quad (22)$$

for  $s \in [(\theta k)^{-1}, 1]$ , all  $\theta \geq 1$ . Therefore, we have just seen that the distribution of deviations between high (large) sample quantiles and their theoretical counterparts is attainable, with a vanishing bias, by means of a different normalization than in the left hand-side of (20). The weak convergence of  $\hat{q}(n/k)$  is supported on the latter.

**Proposition 13** *Let  $X_1, X_2, \dots$  be i.i.d. random variables with tail quantile function  $U$  satisfying condition C in Section 2 (i.e. relation (7)). Suppose  $k = k_n$  is a sequence of positive integers such that  $k_n \rightarrow \infty, k_n/n \rightarrow 0$ , as  $n \rightarrow \infty$ . Then  $\hat{q}(n/k)$  is a consistent estimator for  $q(n/k)$  in the sense that the following convergence in probability holds,*

$$\frac{\hat{q}\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \xrightarrow[n \rightarrow \infty]{p} 1.$$

**Proof:** We begin by noting that

$$\begin{aligned} \frac{\hat{q}\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} &= -\frac{1}{\log 2} \int_0^1 \frac{\widehat{U}\left(\frac{n}{2ks}\right) - \widehat{U}\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \\ &= -\frac{1}{\log 2} \left\{ \int_{\frac{1}{2k}}^1 \frac{X_{n-[2ks],n} - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} - \int_{\frac{1}{k}}^1 \frac{X_{n-[ks],n} - U\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \right. \end{aligned} \quad (23)$$

$$\left. - \int_{\frac{1}{2k}}^{\frac{1}{k}} \frac{X_{n,n} - U\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} + \int_{\frac{1}{2k}}^1 \frac{U\left(\frac{n}{2ks}\right) - U\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \right\}. \quad (24)$$

The two integral terms in (23) shall be handled jointly through the consideration of  $R_2(s)$  (see Eq. (21) with  $\theta = 2$ ) in the one integral below denoted by  $I_1(k, n)$ ,

$$\int_{\frac{1}{2k}}^1 \frac{X_{n-[2ks],n} - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} - \int_{\frac{1}{k}}^1 \frac{X_{n-[ks],n} - U\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} = \int_{\frac{1}{2}}^1 \frac{X_{n-[2ks],n} - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} =: I_1(k, n) \quad (25)$$

as follows:

$$I_1(k, n) = \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \left\{ \int_{\frac{1}{2}}^1 R_2(s) \frac{ds}{s} + \int_{\frac{1}{2}}^1 \left( \frac{a\left(\frac{n}{2ks}\right)}{a\left(\frac{n}{k}\right)} - 1 \right) R_2(s) \frac{ds}{s} \right\}. \quad (26)$$

Now, Lemma 16 ascertains

$$\begin{aligned}
 I_1(k, n) &= \int_{\frac{1}{2}}^1 \frac{X_{n-[2ks],n} - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \\
 &\leq \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \left| \int_{\frac{1}{2}}^1 R_2(s) \frac{ds}{s} \right| + \left( \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \right)^2 \left| \int_{\frac{1}{2}}^1 R_2(s) \log(2s) \frac{ds}{s} \right| \\
 &\leq \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \left( 1 + \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \log 2 \right) \left| \int_{\frac{1}{2}}^1 R_2(s) \frac{ds}{s} \right|, \tag{27}
 \end{aligned}$$

with high probability, for sufficiently large  $n$ . We can provide a similar lower bound.

Owing to (22) the following holds w.r.t. the integral featuring in the upper bound (27), for any positive  $\varepsilon$ ,

$$\left| \int_{\frac{1}{2}}^1 R_2(s) \frac{ds}{s} \right| \leq \left| \frac{1}{\sqrt{2k}} \int_{\frac{1}{2}}^1 s^{-2} W_n(s) ds \right| + \int_{\frac{1}{2}}^1 (s^{-\varepsilon} - 1) \frac{ds}{s} (1 + o_p(1)) - \varepsilon \int_{\frac{1}{2}}^1 s^{-\varepsilon} \log s \frac{ds}{s}.$$

Since  $\varepsilon > 0$  is arbitrary, then

$$0 < \int_{\frac{1}{2}}^1 (s^{-\varepsilon} - 1) \frac{ds}{s} = \frac{2^\varepsilon - 1}{\varepsilon} - \log 2 \xrightarrow{\varepsilon \downarrow 0} 0,$$

meaning that

$$\int_{\frac{1}{2}}^1 (s^{-\varepsilon} - 1) \frac{ds}{s}$$

can be discarded. A similar line of reasoning applies to

$$\varepsilon \int_{\frac{1}{2}}^1 s^{-\varepsilon} \log\left(\frac{1}{s}\right) \frac{ds}{s} = 2^\varepsilon \log 2 - \frac{2^\varepsilon - 1}{\varepsilon} \xrightarrow{\varepsilon \downarrow 0} 0,$$

thus also discarded.

We now recall that  $k = k(n)$  is a sequence of positive integers tending to infinity as  $n \rightarrow \infty$ . Let us define

$$Y_n := \frac{1}{\sqrt{2k}} \int_{\frac{1}{2}}^1 W_n(s) \frac{ds}{s^2},$$

which regards a sequence of normal random variables with zero mean and variance equal to

$$\text{Var}(Y_n) = \frac{1 - \log 2}{k} \xrightarrow{n \rightarrow \infty} 0.$$

The latter means that  $\{Y_n\}_{n \geq 1}$  is a sequence of degenerate random variables, eventually, and the two integrals in (23) (unified in (25); see also Eq. (26) in terms of  $R_2(s)$ ) vanish with probability

tending to one as  $n \rightarrow \infty$ . In this respect we note that  $a(n/k)/q(n/k) = o(1)$ , which entails in fact that

$$I_1(k, n) = o_p(1) \left( = o_p\left(\frac{a(n/k)}{q(n/k)}\right) = O_p\left(\frac{a(n/k)}{\sqrt{k}q(n/k)}\right) \right).$$

The rest of the proof addresses the terms in (24). We note that the first integral in (24),

$$\begin{aligned} I_2(k, n) &:= \int_{\frac{1}{2k}}^{\frac{1}{k}} \frac{X_{n,n} - U\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \\ &= \int_{1/2}^1 \frac{X_{n,n} - U\left(\frac{n}{s}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \\ &\stackrel{d}{=} \frac{a(n)}{q\left(\frac{n}{k}\right)} \left\{ \frac{U\left(\frac{1}{U_{1,n}}\right) - U(n)}{a(n)} \log 2 - \int_{\frac{1}{2}}^1 \frac{U\left(\frac{n}{s}\right) - U(n)}{a(n)} \frac{ds}{s} \right\} \\ &= \frac{a(n)}{q\left(\frac{n}{k}\right)} \left\{ -\log(n U_{1,n}) \log 2 + \int_{\frac{1}{2}}^1 \log s \frac{ds}{s} \right\} (1 + o_p(1)) \\ &= \frac{a(n)}{q\left(\frac{n}{k}\right)} \log 2 \left( -\log(n U_{1,n}) - \frac{1}{2} \log 2 \right) (1 + o_p(1)). \end{aligned}$$

Now, the probability integral transformation yields the following equality in distribution for the random term above:

$$-\log(n U_{1,n}) \stackrel{d}{=} E_{n,n} - \log n, \quad (28)$$

where  $E_{n,n}$  is the maximum of  $n$  i.i.d. standard exponential random variables. Hence, the random variable (28) converges in distribution to a Gumbel random variable with distribution function given by  $\exp\{-e^{-x}\}$ ,  $x \in \mathbb{R}$ . Moreover,  $a(n)/q(n/k) \rightarrow 0$ , as  $n \rightarrow \infty$ , because  $a(n/k)/q(n/k) = o(1)$  (see Lemma 15(2) in Appendix C), where the auxiliary positive function  $a$  satisfies  $a(t) \rightarrow 0$ , as  $t \rightarrow \infty$ , by assumption. Therefore,

$$I_2(k, n) = o_p(1) \left( = O_p\left(\frac{a(n)}{q(n/k)}\right) \right). \quad (29)$$

In order to finally attain consistency of  $\hat{q}(n/k)$  let us consider the last integral in (24), which we will show it is bounded. We establish the upper bound,

$$\int_{\frac{1}{2k}}^1 \frac{U\left(\frac{n}{ks}\right) - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \leq \int_0^1 \frac{U\left(\frac{n}{ks}\right) - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \quad (30)$$

and the lower bound,

$$\begin{aligned} & \int_{\frac{1}{2k}}^1 \frac{U\left(\frac{n}{ks}\right) - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \\ &= \int_0^{1-\frac{1}{2k}} \frac{U\left(\frac{n}{ks+1/2}\right) - U\left(\frac{n}{2ks+1}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s + \frac{1}{2k}} \\ &\geq \int_0^1 \frac{U\left(\frac{n}{ks+1/2}\right) - U\left(\frac{n}{2ks+1}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s + \frac{1}{2k}} - 2 \int_{1-\frac{1}{2k}}^1 \frac{U\left(\frac{n}{ks+1/2}\right) - U\left(\frac{n}{2ks+1}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s + \frac{1}{2k}}. \end{aligned}$$

Making  $t = n/k$  run on the real line towards infinity, then the following condition of  $\Pi$ -variation (i.e. condition (10)),

$$\lim_{t \rightarrow \infty} \frac{\int_0^1 U\left(\frac{tx}{s}\right) \frac{ds}{s} - \int_0^1 U\left(\frac{t}{s}\right) \frac{ds}{s}}{q(t)} = \log x, \quad x > 0, \quad (31)$$

clearly entails the following limit for the upper bound in (30):

$$\frac{\int_0^1 U\left(\frac{n}{ks}\right) \frac{ds}{s} - \int_0^1 U\left(\frac{n}{2ks}\right) \frac{ds}{s}}{q\left(\frac{n}{k}\right)} = - \frac{\int_0^1 U\left(\frac{n}{2ks}\right) \frac{ds}{s} - \int_0^1 U\left(\frac{n}{ks}\right) \frac{ds}{s}}{q\left(\frac{n}{k}\right)} \xrightarrow{n \rightarrow \infty} \log 2.$$

Regarding the lower bound provided above, that is,

$$\int_{\frac{1}{2k}}^1 \frac{U\left(\frac{n}{ks}\right) - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \geq \int_0^1 \frac{U\left(\frac{n}{k\left(s+\frac{1}{2k}\right)}\right) - U\left(\frac{n}{2k\left(s+\frac{1}{2k}\right)}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s + \frac{1}{2k}} \quad (32)$$

$$- 2 \frac{q\left(\frac{n}{2k}\right)}{q\left(\frac{n}{k}\right)} \int_{1-\frac{1}{2k}}^1 \frac{U\left(\frac{n}{k\left(s+\frac{1}{2k}\right)}\right) - U\left(\frac{n}{2k\left(s+\frac{1}{2k}\right)}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s + \frac{1}{2k}}, \quad (33)$$

we note that for every  $\varepsilon > 0$ , there exists  $n_0 \in \mathbb{N}$  such that for  $n \geq n_0$ ,

$$\left| \frac{1}{s + 1/(2k)} - \frac{1}{s} \right| < \varepsilon. \quad (34)$$

Whence, we have in turn the following inequality with respect to (32):

$$\int_0^1 \frac{U\left(\frac{n}{k\left(s+\frac{1}{2k}\right)}\right) - U\left(\frac{n}{2k\left(s+\frac{1}{2k}\right)}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s + \frac{1}{2k}} > \int_0^1 \frac{U\left(\frac{n}{k\left(s+\frac{1}{2k}\right)}\right) - U\left(\frac{n}{2k\left(s+\frac{1}{2k}\right)}\right)}{q\left(\frac{n}{k}\right)} \left(\frac{1}{s} - \varepsilon\right) ds.$$

For the first part of the right-hand side of the above we use again condition (31), while the second



part is dealt with Theorem B.2.19 of de Haan and Ferreira (2006) involving the fact that  $U \in \Pi(a)$ :

$$\begin{aligned} & \int_0^1 \frac{U\left(\frac{n}{k\left(s+\frac{1}{2k}\right)}\right) - U\left(\frac{n}{2k\left(s+\frac{1}{2k}\right)}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} + \varepsilon \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \int_0^1 \frac{U\left(\frac{n}{2k\left(s+\frac{1}{2k}\right)}\right) - U\left(\frac{n}{k\left(s+\frac{1}{2k}\right)}\right)}{a\left(\frac{n}{k}\right)} ds \\ &= \log 2(1 + o(1)) - \varepsilon \frac{a\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} \log 2(1 + o(1)) \\ &\xrightarrow{n \rightarrow \infty} \log 2. \end{aligned}$$

For the latter, we recall that  $a(n/k) = o(q(n/k))$ .

Now we write  $\delta = 1/(2k) > 0$  everywhere in (33). Furthermore, we assume that there exists  $n_0 \in \mathbb{N}$  such that, for  $n \geq n_0$ , the term  $n\delta$  is large enough and the integral in (33) can be rephrased as

$$I_\delta^* := \frac{\int_{1-\delta}^1 \left( U\left(\frac{2}{s+\delta}n\delta\right) - U\left(\frac{1}{s+\delta}n\delta\right) \right) \frac{ds}{s+\delta}}{\int_{n\delta}^1 a(s) \frac{ds}{s}}. \quad (35)$$

We note that, for every fixed  $\delta > 0$ , we have that from the  $\Pi$ -variation of  $U$  that the following holds for the numerator of  $I_\delta^*$  properly rescaled by  $a(n\delta)$  (cf. Theorem B.2.19 in de Haan and Ferreira, 2006):

$$\frac{\int_{1-\delta}^1 \left( U\left(\frac{2}{s+\delta}n\delta\right) - U\left(\frac{1}{s+\delta}n\delta\right) \right) \frac{ds}{s+\delta}}{a(n\delta)} \xrightarrow{n \rightarrow \infty} \int_{1-\delta}^1 \log 2 \frac{ds}{s+\delta} = \log(1+\delta) \log 2.$$

For arbitrary small  $\delta$ , the latter approaches zero. Predicated on the above, we apply Cauchy's rule to obtain  $\lim_{\delta \rightarrow 0} I_\delta^*$  (we recall that  $\delta \rightarrow 0$  implies  $n \rightarrow \infty$ ). Towards this end, we apply Eq. (2.11) of Chiang (2000) upon the numerator of  $I_\delta^*$ , whence

$$\begin{aligned} \lim_{\delta \rightarrow 0} I_\delta^* &= \lim_{\delta \rightarrow 0} \frac{\int_{1-\delta}^1 \left( U'\left(\frac{2}{s+\delta}n\delta\right) \frac{2s}{(s+\delta)^3} - U'\left(\frac{1}{s+\delta}n\delta\right) \frac{s}{(s+\delta)^3} \right) ds}{-\frac{a(n\delta)}{n\delta}} \\ &\quad + \lim_{\delta \rightarrow 0} \left\{ \delta \int_{1-\delta}^1 \frac{U\left(\frac{2n\delta}{s+\delta}\right) - U\left(\frac{n\delta}{s+\delta}\right)}{a(n\delta)} \frac{ds}{(s+\delta)^2} - \delta \frac{U(2n\delta) - U(n\delta)}{a(n\delta)} \right\}. \end{aligned}$$

Since  $U'(t) = a(t)/t$  then the limit becomes equal to the the limit of

$$-\int_{1-\delta}^1 \left( \frac{a\left(\frac{2n\delta}{s+\delta}\right)}{a(n\delta)} - \frac{a\left(\frac{n\delta}{s+\delta}\right)}{a(n\delta)} \right) \frac{s ds}{(s+\delta)^2} + \delta \left( \int_{1-\delta}^1 \frac{U\left(\frac{2n\delta}{s+\delta}\right) - U\left(\frac{n\delta}{s+\delta}\right)}{a(n\delta)} \frac{ds}{(s+\delta)^2} - \frac{U(2n\delta) - U(n\delta)}{a(n\delta)} \right).$$

We can now take any arbitrary small  $\delta$  (making  $n \rightarrow \infty$ ) in order to apply the uniform convergence of  $a \in RV_0$  and  $U \in \Pi(a)$  so that the above integrals are ensured finite and then equal to zero by definition. Hence, all the terms are negligible as  $\delta$  converges to zero meaning that  $\lim_{\delta \rightarrow 0} I_\delta^*$

becomes null. Therefore,

$$\int_{\frac{1}{2k}}^1 \frac{U\left(\frac{n}{2ks}\right) - U\left(\frac{n}{ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} \xrightarrow{n \rightarrow \infty} -\log 2.$$

and the precise result for consistency of  $\hat{q}(n/k)$  thus follows by noting that  $q(n/k) \sim q(n/(2k))$ .  $\square$

## B Asymptotic distribution of $\hat{q}(n/k)$

In order to establish the asymptotic distribution of the proposed estimator for  $q(n/k)$  (see (12) or (13)), we need to gain further insight about the distributional representation obtained from (20). Therefore, the proper framework calls for a second order refinement of (6). Specifically, if the tail quantile function satisfies the second order condition (15), then Theorem 2.4.2 of de Haan and Ferreira (2006) ascertains that, for each  $\varepsilon > 0$ ,

$$\sup_{\frac{1}{\theta k} \leq s \leq 1} s^{1/2+\varepsilon} \left| \sqrt{\theta k} \left( \frac{X_{n-[\theta ks],n} - U\left(\frac{n}{\theta k}\right)}{a_0\left(\frac{n}{\theta k}\right)} + \log s \right) - \frac{W_n(s)}{s} - \sqrt{\theta k} A_0\left(\frac{n}{\theta k}\right) \frac{1}{2} (\log s)^2 \right| \xrightarrow[n \rightarrow \infty]{p} 0, \quad (36)$$

provided  $k = k_n \rightarrow \infty$ ,  $k_n/n = o(n)$  and  $\sqrt{k_n} A_0(n/k_n) = O(1)$ .

Therefore, the asymptotic distribution of  $\hat{q}(n/k)$  will appear intertwined with the proof of consistency in Proposition 13 via  $R_\theta(s)$ , (defined in (21) for  $s \in [(\theta k)^{-1}, 1]$ , see also (22)), albeit under the second order grasp provided above. Hence, we have the following Proposition (cf. (2.4.7) of de Haan and Ferreira, 2006).

**Proposition 14** *Given condition C from Section 2, suppose the second order condition (15) holds. Let  $k = k_n \rightarrow \infty$ ,  $k_n/n = o(n)$  and  $\sqrt{k_n} A(n/k_n) \rightarrow \lambda \in \mathbb{R}$ , as  $n \rightarrow \infty$ . Then, for  $\theta \geq 1$  and for each  $\varepsilon > 0$  sufficiently small,*

$$\sup_{\frac{1}{\theta k} \leq s \leq 1} s^{1/2+\varepsilon} \left| \sqrt{\theta k} \frac{X_{n-[\theta ks],n} - U\left(\frac{n}{\theta ks}\right)}{a\left(\frac{n}{\theta ks}\right)} - \frac{W_n(s)}{s} \right| = o_p(1).$$

**Proof:** Similarly to the equality right after (21), we have that

$$R_\theta(s) := \frac{X_{n-[\theta ks],n} - U\left(\frac{n}{\theta ks}\right)}{a\left(\frac{n}{\theta ks}\right)} = \frac{a_0\left(\frac{n}{\theta k}\right)}{a\left(\frac{n}{\theta ks}\right)} \left\{ \frac{X_{n-[\theta ks],n} - U\left(\frac{n}{\theta k}\right)}{a_0\left(\frac{n}{\theta k}\right)} - \frac{U\left(\frac{n}{\theta ks}\right) - U\left(\frac{n}{\theta k}\right)}{a_0\left(\frac{n}{\theta k}\right)} \right\}.$$

Noting that

$$\frac{a_0(t)}{a\left(\frac{t}{s}\right)} = \frac{a_0(t)}{a(t)} \frac{a(t)}{a\left(\frac{t}{s}\right)},$$

for all  $s > 0$ , then Lemma 16 combined with Remark 17 yields the expansion

$$\frac{a_0(t)}{a(\frac{t}{s})} = \frac{a_0(t)}{a(t)} \left( 1 - \frac{a(t)}{q(t)} \log s + o\left(\frac{a(t)}{q(t)}\right) \right) = \frac{a_0(t)}{a(t)} \left( 1 + A(t) \log s + o(A(t)) \right), \quad (37)$$

for all  $s > 0$ . In this respect, we also note that  $|A| \in RV_0$  and  $a_0(t)/a(t) = 1 + o(A(t))$ .

Having set  $1/(\theta k) \leq s \leq 1$ , we thus have from (36), the uniform bounds in (40) and the second equality in (37), that

$$\begin{aligned} \sqrt{\theta k} R_\theta(s) &= \frac{W_n(s)}{s} + A\left(\frac{n}{\theta k}\right) \frac{\log s}{s} W_n(s) \mp \varepsilon s^{-\varepsilon} \sqrt{\theta k} A\left(\frac{n}{\theta k}\right) \pm \varepsilon s^{-\varepsilon} \log s \sqrt{\theta k} A^2\left(\frac{n}{\theta k}\right) \\ &\quad + o_p(s^{-\frac{1}{2}-\varepsilon}) + o_p\left(s^{-\frac{1}{2}-\varepsilon} \log s A\left(\frac{n}{\theta k}\right)\right), \end{aligned}$$

uniformly in  $s$ . Hence, the assumption that  $\sqrt{k}A(n/k) = O(1)$  entails that  $\log(1/s)A(n/(\theta k)) \rightarrow 0$ , whereas  $\varepsilon s^{-\varepsilon} \sqrt{\theta k} A(n/(\theta k))$  virtually becomes  $o(s^{-1/2-\varepsilon})$  for each  $\varepsilon > 0$  arbitrarily small and uniformly in  $s \in [(\theta k)^{-1}, 1]$ . The  $o_p$ -terms are uniform in  $s \in [1/(\theta k), 1]$ . Hence the following representation for  $\sqrt{\theta k} R_\theta(s)$ , valid for  $\varepsilon \in (0, 1)$ ,

$$\sqrt{\theta k} R_\theta(s) = \frac{W_n(s)}{s} + o_p(s^{-1/2-\varepsilon}).$$

□

**Proof of Theorem 8:** Similarly as in (25), we have that

$$\begin{aligned} &\frac{q\left(\frac{n}{k}\right)}{a\left(\frac{n}{k}\right)} \left( \frac{\hat{q}\left(\frac{n}{k}\right)}{q\left(\frac{n}{k}\right)} - 1 \right) \\ &= -\frac{1}{a\left(\frac{n}{k}\right)} \frac{1}{\log 2} \left\{ \int_{\frac{1}{2}}^1 \left( X_{n-[2ks],n} - U\left(\frac{n}{2ks}\right) \right) \frac{ds}{s} - \int_{\frac{1}{2k}}^{\frac{1}{k}} \left( X_{n,n} - U\left(\frac{n}{ks}\right) \right) \frac{ds}{s} \right. \\ &\quad \left. - q\left(\frac{n}{k}\right) \left( \int_{\frac{1}{2k}}^1 \frac{U\left(\frac{n}{ks}\right) - U\left(\frac{n}{2ks}\right)}{q\left(\frac{n}{k}\right)} \frac{ds}{s} - \log 2 \right) \right\} \\ &= -\frac{1}{\log 2} \{ J_1(k, n) - J_2(k, n) \} + \frac{q\left(\frac{n}{k}\right)}{a\left(\frac{n}{k}\right)} \frac{1}{\log 2} J_3(k, n). \end{aligned} \quad (38)$$

By mimicking the steps of progression from (25) to (26), we obtain for the first integral above that

$$\begin{aligned}\sqrt{2k} J_1(k, n) &:= \sqrt{2k} \int_{\frac{1}{2}}^1 \frac{X_{n-[2ks],n} - U\left(\frac{n}{2ks}\right)}{a\left(\frac{n}{k}\right)} \frac{ds}{s} \\ &= \int_{\frac{1}{2}}^1 \sqrt{2k} R_2(s) \frac{ds}{s} + \int_{\frac{1}{2}}^1 \left( \frac{a\left(\frac{n}{2ks}\right)}{a\left(\frac{n}{k}\right)} - 1 \right) \sqrt{2k} R_2(s) \frac{ds}{s}.\end{aligned}$$

Hence, Proposition 14 while assuming that  $\sqrt{k} a(n/k)/q(n/k) = O(1)$  (by appointment of Remark 17) and application of the uniform bounds in (41) with  $a_0(t) := a(t)(1+o(A(t)))$  and  $A_0(t) := A(t)$ , imply for each  $\varepsilon > 0$ ,

$$\sqrt{2k} J_1(k, n) = \int_{\frac{1}{2}}^1 W_n(s)(1 - \log(2s)) \frac{ds}{s^2} + o_p(1) \int_{\frac{1}{2}}^1 \log\left(\frac{1}{2s}\right) \left(\frac{1}{s}\right)^{3/2+\varepsilon} ds + o_p\left(A_0\left(\frac{n}{k}\right)\right).$$

Since the integral  $\int_{1/2}^1 W_n(s)(1 - \log(2s))s^{-2} ds$  converges to a sum of independent normal random variables, then the expression above allows to conclude that the first random component in (38) is negligible with high probability because

$$J_1(k, n) = O_p\left(\frac{1}{\sqrt{k}}\right).$$

Now, similarly to  $I_2(k, n)$  in the proof of Proposition 13, albeit under the second order condition (15) and pertaining uniform bounds provided by (40), we now have that

$$\begin{aligned}J_2(k, n) &:= \int_{\frac{1}{2k}}^{\frac{1}{k}} \frac{X_{n,n} - U\left(\frac{n}{ks}\right)}{a\left(\frac{n}{k}\right)} \frac{ds}{s} \\ &= \frac{a(n)}{a\left(\frac{n}{k}\right)} \left\{ -\log 2 \log(nU_{1,n}) + \frac{a_0(n)}{a(n)} \int_{\frac{1}{2}}^1 \log s \frac{ds}{s} + A_0(n) \int_{\frac{1}{2}}^1 \left( \frac{(\log s)^2}{2} \pm \varepsilon s^{-\varepsilon} \right) \frac{ds}{s} \right\}.\end{aligned}$$

Again, note that  $a_0(n)/a(n) - 1 = o(A(n))$  and  $A_0(n) = A(n)$ . Hence,

$$\begin{aligned}\frac{a\left(\frac{n}{k}\right)}{a(n)} \frac{1}{\log 2} J_2(k, n) &= -\log(nU_{1,n}) - \frac{\log 2}{2} + \frac{1}{\log 2} A(n) \int_{\frac{1}{2}}^1 \left( \frac{(\log s)^2}{2} \pm \varepsilon s^{-\varepsilon} \right) \frac{ds}{s} + o(A(n)) \\ &= -\log(nU_{1,n}) - \frac{\log 2}{2} + o(1).\end{aligned}$$

Furthermore, assuming that  $k = k(n)$  is such that  $a(n)/a(n/k) \rightarrow 1$ , then the following convergence in distribution holds

$$\frac{a\left(\frac{n}{k}\right)}{a(n)} \frac{1}{\log 2} J_2(k, n) \xrightarrow[n \rightarrow \infty]{d} \Lambda - \frac{\log 2}{2},$$

where  $\Lambda$  denotes a Gumbel random variable with distribution function  $\exp\{-e^{-x}\}$ ,  $x \in \mathbb{R}$  (cf. Eq. (28) and subsequent text). We note at this point that, if  $a(n)/a(n/k)$  converges to a constant different than 1, then a change in the scale is performed. The following also holds provided (41) and that  $\sqrt{k_n} A(n/k_n) = O(1)$ :

$$\frac{1}{\log 2} J_2(k, n) \xrightarrow[n \rightarrow \infty]{d} \Lambda - \frac{\log 2}{2},$$

Finally we turn to the bias term  $J_3(k, n)$ . By assumption,

$$\frac{J_3(k, n)}{A(\frac{n}{k})} = \frac{1}{A(\frac{n}{k})} \left( \int_{\frac{1}{2k}}^1 \frac{U(\frac{n}{ks}) - U(\frac{n}{2ks})}{q(\frac{n}{k})} \frac{ds}{s} - \log 2 \right) \xrightarrow[n \rightarrow \infty]{} \lambda,$$

as  $n \rightarrow \infty$ . Therefore, since  $A(\frac{n}{k}) \sim -a(\frac{n}{k})/q(\frac{n}{k})$  (cf. Remark 17), the deterministic term  $J_3(k, n)$  renders the following contribution to the asymptotic bias:

$$\frac{q(\frac{n}{k})}{a(\frac{n}{k})} \frac{1}{\log 2} J_3(k, n) \xrightarrow[n \rightarrow \infty]{} -\frac{\lambda}{\log 2}.$$

□

## C Auxiliary Results

### Lemma 15

1. Suppose  $U \in \Pi(a)$ . Then,

- (i) there exists a positive function  $a$  satisfying  $a(t) \sim a_0(t)$ , as  $t \rightarrow \infty$ , such that for any  $\varepsilon > 0$  there exists  $t_0 = t_0(\varepsilon)$  such that, for  $t, st \geq t_0$ ,  $s \in (0, 1]$ ,

$$\left| \frac{U(st) - U(t)}{a(t)} - \log s \right| \leq \varepsilon \max(s^\varepsilon, s^{-\varepsilon});$$

- (ii)  $a \in RV_0$  and for any  $\varepsilon > 0$  there exists  $t_0 = t_0(\varepsilon)$  such that, for  $t, st \geq t_0$ ,  $s \in (0, 1]$ ,

$$\left| \frac{a(st)}{a(t)} - 1 \right| \leq \varepsilon \max(s^\varepsilon, s^{-\varepsilon}).$$

2. Suppose  $a > 0$  is a slowly varying function, integrable over finite intervals of  $\mathbb{R}^+$  such that

$$\int_t^\infty a(s) \frac{ds}{s} < \infty.$$

for every  $t > 0$ . Then  $a(t) \rightarrow 0$ , as  $t \rightarrow \infty$ , and

$$\lim_{t \rightarrow \infty} \int_t^\infty \frac{a(s)}{a(t)} \frac{ds}{s} = \infty.$$

**Proof:** Part 1.(i) of the Lemma comes from de Haan and Ferreira (2006, cf. Proposition B.2.17), which is closely followed by (ii), a result from Drees (1998) (cf. Proposition B.1.10 of de Haan and Ferreira, 2006). The second part follows from Karamata's theorem for regularly varying functions (cf. Theorem B.1.5 of de Haan and Ferreira, 2006).  $\square$

The following lemma regards a second order condition on the auxiliary function  $a$ :

**Lemma 16** Under condition C (i.e. relation (7) provided in Section 2, the following limit holds with  $q(t) := \int_t^\infty a(s) ds/s$  (defined in (8)):

$$\lim_{t \rightarrow \infty} \frac{\frac{a(tx)}{a(t)} - 1}{\frac{a(t)}{q(t)}} = -\log x, \quad x > 0.$$

**Proof:** The underlying assumption that  $U \in \Pi(a)$  entails

$$\begin{aligned} \frac{q(t)}{a(t)} \left( \frac{a(tx)}{a(t)} - 1 \right) &= \frac{q(t)}{U(tx) - U(t)} \frac{U(tx) - U(t)}{a(t)} \left( \frac{a(tx)}{a(t)} - 1 \right) \\ &= \frac{q(t)}{U(tx) - U(t)} \log x \left( \frac{a(tx)}{a(t)} - 1 \right) (1 + o(1)). \quad (t \rightarrow \infty) \end{aligned} \quad (39)$$

Furthermore, according to the main relation (7),

$$\frac{q(t)}{U(tx) - U(t)} = \frac{\int_t^\infty a(s) \frac{ds}{s}}{\int_t^{tx} a(s) \frac{ds}{s} (1 + o(1))} = 1 + \frac{\int_{tx}^\infty a(s) \frac{ds}{s}}{\int_t^{tx} a(s) \frac{ds}{s}} (1 + o(1)).$$

By taking the limit of the latter term when  $t \rightarrow \infty$ , we get from Cauchy's rule together with the fundamental theorem of integral calculus that

$$\lim_{t \rightarrow \infty} \frac{\int_{tx}^\infty a(s) \frac{ds}{s}}{\int_t^{tx} a(s) \frac{ds}{s}} = \lim_{t \rightarrow \infty} \frac{-a(tx)}{a(tx) - a(t)} = -\lim_{t \rightarrow \infty} \left( \frac{a(tx)}{a(t)} - 1 \right)^{-1}.$$

Giving heed to (39), the limiting statement follows in a straightforward manner:

$$\begin{aligned} \frac{q(t)}{a(t)} \left( \frac{a(tx)}{a(t)} - 1 \right) &= \log x \left( 1 + \frac{\int_{tx}^{\infty} a(s) \frac{ds}{s}}{\int_t^{tx} a(s) \frac{ds}{s}} \right) \left( \frac{a(tx)}{a(t)} - 1 \right) (1 + o(1)) \\ &= -\log x + \log x \left( \frac{a(tx)}{a(t)} - 1 \right) (1 + o(1)). \quad (t \rightarrow \infty) \end{aligned}$$

□

In addition to the second order condition (15), Theorem 2.3.6 of de Haan and Ferreira (2006) ascertains the existence of functions  $a_0$  and  $A_0$  satisfying, as  $t \rightarrow \infty$ ,  $A_0(t) \sim A(t)$  and  $a_0(t)/a(t) - 1 = o(A(t))$ , with the property that for any  $\varepsilon > 0$ , there exists  $t_0 = t_0(\varepsilon)$  such that for all  $t, tx \geq t_0$ ,

$$\left| \frac{\frac{U(tx) - U(t)}{a_0(t)} - \log x}{A_0(t)} - \frac{1}{2}(\log x)^2 \right| \leq \varepsilon \max(x^\varepsilon, x^{-\varepsilon}) \quad (40)$$

and

$$\left| \frac{\frac{a_0(tx)}{a_0(t)} - 1}{A_0(t)} - \log x \right| \leq \varepsilon \max(x^\varepsilon, x^{-\varepsilon}). \quad (41)$$

**Remark 17** We note that relation (41) combined with Lemma 16 ascertains that  $-a_0(t)/q(t) = cA_0(t)$ , with  $c \neq 0$  because  $\rho = \gamma = 0$  (cf. Eq. (B.3.4) and Remark B.3.5 in de Haan and Ferreira, 2006). Hence the assumption in this paper that the function  $q$  can be redefined in order that  $-a/q \sim A$  is satisfied.

## References

- Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probab.*, 2:792–804.
- Cai, J. J., de Haan, L., and Zhou, C. (2012). Bias correction in extreme value statistics with index around zero. *Extremes*, 25:DOI 10.1007/s10687-012-0158-x.
- Chiang, A. C. (2000). *Elements of Dynamic Optimization*. Waveland Press.
- de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes*. Mathematisch Centrum Amsterdam.
- de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer.

- Dekkers, A., Einmahl, J., and de Haan, L. (1989). A moment estimator for the index of an extremevalue distribution. *Ann. Statist.*, 17:1833–1855.
- Drees, H. (1998). On smooth statistical tail functionals. *Scand. J. Statist.*, 25:187–210.
- Einmahl, J. H. J. and Magnus, J. R. (2008). Records in Athletics through Extreme-Value Theory. *J. Amer. Statist. Assoc.*, 103:1382–1391.
- Falk, M. (1995). Some best parameter estimates for distributions with finite endpoint. *Statistics*, 27:115–125.
- Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Cambridge Philosophical Society. Mathematical Proceedings*, 24:180–190.
- Fraga Alves, I., de Haan, L., and Neves, C. (2013). How far can Man go? In Torelli, N., Pesarin, F., and Bar-Hen, A., editors, *Advances in Theoretical and Applied Statistics*, pages 187–197. Springer Berlin Heidelberg.
- Gardes, L., Girard, S., and Guillou, A. (2011). Weibull tail-distributions revisited: A new look at some tail estimators. *J. Statist. Plann. Inference*, 141:429–444.
- Girard, S., Guillou, A., and Stupfler, G. (2012). Estimating an endpoint with high-order moments. *TEST*, 21:697–729.
- Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d’une série aléatoire. *Anal. Math.*, 44:423–453.
- Goegebeur, Y., Beirlant, J., and De Wet, T. (2010). Generalized kernel estimators for the Weibull-tail coefficient. *Comm. Statist. Theory Methods*, 39:3695–3716.
- Hall, P. (1982). On estimating the endpoint of a distribution. *Ann. Statist.*, 10:556–568.
- Hall, P. and Wang, J. Z. (1999). Estimating the end-point of a probability distribution using minimum-distance methods. *Bernoulli*, 5:177–189.
- Li, D. and Peng, L. (2009). Does bias reduction with external estimator of second order parameter work for endpoint? *J. Statist. Plann. Inference*, 139:1937–1952.
- Li, D., Peng, L., and Xu, X. (2011). Bias reduction for endpoint estimation. *Extremes*, 14:393–412.



- Li, Z. and Peng, L. (2012). Bootstrapping endpoint. *Sankhyā*, 74:126–140.
- Neves, C. and Pereira, A. (2010). Detecting finiteness in the right endpoint of light-tailed distributions. *Statist. Probab. Lett.*, 80:437–444.
- Scholz, F. W. (2003). Statistical extreme value analysis of anc taxiway centerline deviations for 747 aircraft. *FAA/Boeing Cooperative Research and Development Agreement 01-CRDA-0164*. [http://www.faa.gov/airports/resources/publications/reports/media/ANC\\_747.pdf](http://www.faa.gov/airports/resources/publications/reports/media/ANC_747.pdf).