

On the block maxima method in extreme value theory

Ana Ferreira

*ISA, Univ Tecn Lisboa, Tapada da Ajuda 1349-017 Lisboa, Portugal
CEAUL, FCUL, Bloco C6 - Piso 4 Campo Grande, 749-016 Lisboa, Portugal
anafh@isa.utl.pt*

and

Laurens de Haan

*Erasmus Univ Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
CEAUL, FCUL, Bloco C6 - Piso 4 Campo Grande, 749-016 Lisboa, Portugal
ldehaan@ese.eur.nl*

Abstract: In extreme value theory there are two fundamental approaches, both widely used: the block maxima method and the peaks-over-threshold (POT) method. Whereas much research has gone into the POT method, the block maxima method has not been studied thoroughly. The present paper aims at providing conditions under which the block maxima method can be justified. In this paper we restrict attention to the independent and identically distributed case and focus on the probability weighted moment (PWM) estimators of Hosking, Wallis and Wood (1985).

MSC 2010 subject classifications: Primary 62G32, ; secondary 62G20, 62G30.

Keywords and phrases: block maxima, probability weighted moment estimators, extreme value index, asymptotic normality, extreme quantile estimation.

1. Introduction

The *block maxima* approach in extreme value theory (EVT), consists of dividing the observation period into non-overlapping periods of equal size and restricts attention to the maximum observation in each period. The new observations thus created follow - under extreme value conditions - approximately an extreme value distribution (G_γ for some real γ). Parametric statistical methods for the extreme value distributions are then applied to those observations. Usually it is taken for granted that the block maxima follow very well an extreme value distribution. In this paper we take this misspecification into account. Since G_γ is not the exact distribution for those observations, a bias may appear. The procedure can be justified using (a strengthening of) the domain of attraction conditions of EVT.

In the *peaks-over-threshold* approach in EVT one selects those of the initial observations that exceed a certain high threshold. The probability distribution of those selected observations - under extreme value conditions - is approximately a generalized Pareto distribution (GPD). Parametric statistical methods for GPD

are then applied to those observations. Again, a bias may appear since GPD is not the exact distribution of those selected observations.

The block maxima method is the older one (see e.g. Gumbel, 1958). The POT method has been developed by Pickands (1975) who provided the theoretical framework and devised statistical tools.

In the case of the POT method, exact conditions under which the method is justified are well-known (see e.g. de Haan and Ferreira 2006, Chapters 3-4, and Drees 1998).

The POT method picks up all relevant high observations. The block maxima method, on the one hand misses some of these high observations and, on the other hand, retains some lower observations. Hence the POT seems to make better use of the available information. However, there may be reasons for using the block maxima method:

- The only available information may be block maxima (e.g. yearly maxima).
- The block maxima method may be preferable when the observations are not exactly independent and identically distributed (i.i.d.). For example, there may be a seasonal periodicity in case of yearly maxima or, there may be short range dependence that plays a role within blocks but not between blocks.
- The block maxima method may be easier to apply since the block periods appear naturally in many situations. Hence the problem of choosing a high threshold in the POT method (which is a difficult one) does not play a role.

The present paper aims at formulating exact conditions under which the block maxima method can be justified. Since some of the block maxima may actually not be very high, one expects somewhat more strict conditions in this case than in the POT case. However, as it turns out, the conditions are similar.

Throughout the paper we assume that the observations are i.i.d. When working with block maxima there are two major sets of estimators that are widely used: the maximum likelihood estimators (e.g. Prescott and Walden, 1980) and the probability weighted moment (PWM) estimators (Hosking, Wallis and Wood, 1985). Recently, Dombry (2013) has proved consistency of the maximum likelihood estimators. The present paper concentrates on the PWM estimators.

The asymptotic normality result for the PWM estimators is stated in Section 2, along with a similar result for the accompanying high quantile estimator. The proofs (in Section 3) are based on a uniform expansion of the relevant quantile process given in Proposition 3.1.

In future work we shall extend the results to the non - i.i.d. case and to the maximum likelihood estimator. We shall also develop a theoretical comparison between the peaks-over-threshold and the block maxima methods. A comparative simulation study has been carried out by S. Caires (2009).

2. The estimators and their properties

Let $\tilde{X}_1, \tilde{X}_2, \dots$ be i.i.d. random variables with distribution function F . Define for $m = 1, 2, \dots$ and $i = 1, 2, \dots, k$ the block maxima

$$X_i = \max_{(i-1)m < j \leq im} \tilde{X}_j. \quad (1)$$

Hence, the $m \times k$ observations are divided into k blocks of size m . Write $n = m \times k$, the total number of observations. We study the model for large k and m , hence we shall assume that $n \rightarrow \infty$; in order to obtain meaningful limit results, we have to require that both $m = m_n \rightarrow \infty$ and $k = k_n \rightarrow \infty$, as $n \rightarrow \infty$.

The main assumption is that F is in the domain of attraction of some extreme value distribution

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

Then X_i , when normalized, will follow approximately this extreme value distribution i.e., for appropriately chosen $a_m > 0$ and b_m and all x

$$\lim_{m \rightarrow \infty} P\left(\frac{X_i - b_m}{a_m} \leq x\right) = \lim_{m \rightarrow \infty} F^m(a_m x + b_m) = G_\gamma(x), \quad i = 1, 2, \dots, k. \quad (2)$$

This can be written as

$$\lim_{m \rightarrow \infty} \frac{1}{m - \log F(a_m x + b_m)} = (1 + \gamma x)^{1/\gamma},$$

which is equivalent to the convergence of the inverse functions:

$$\lim_{m \rightarrow \infty} \frac{V(mx) - b_m}{a_m} = \frac{x^\gamma - 1}{\gamma}, \quad x > 0,$$

with $V = (-1/\log F)^\leftarrow$. Hence b_m can be chosen to be $V(m)$. This is the first order condition. For our analysis we also need a second order expansion as follows.

Condition 2.1 (Second order condition). *Suppose that for some positive function a and some positive or negative function A with $\lim_{t \rightarrow \infty} A(t) = 0$,*

$$\lim_{t \rightarrow \infty} \frac{\frac{V(tx) - V(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} du ds = H_{\gamma, \rho}(x),$$

for all $x > 0$ (see e.g. de Haan and Ferreira, Corollary 2.3.4). Note that the function $|A|$ is regularly varying with index $\rho \leq 0$.

The PWM estimators to estimate γ , as well as the location b_m and scale a_m , are defined as follows. Let $X_{1,k}, \dots, X_{k,k}$ be the order statistics of X_1, \dots, X_k . Let

$$M_r = \frac{1}{k} \sum_{i=1}^k \frac{(i-1) \dots (i-r)}{(k-1) \dots (k-r)} X_{i,k}, \quad \text{for } r = 0, 1, 2; k > r. \quad (3)$$

The PWM estimators are simple functionals of M_0 , M_1 and M_2 . The estimator $\hat{\gamma}_{k,m}$ for γ is defined as the solution of the equation

$$\frac{3^{\hat{\gamma}_{k,m}} - 1}{2^{\hat{\gamma}_{k,m}} - 1} = \frac{3M_2 - M_0}{2M_1 - M_0}. \quad (4)$$

The estimator $\hat{a}_{k,m}$ of a_m is

$$\hat{a}_{k,m} = \frac{\hat{\gamma}_{k,m}}{2^{\hat{\gamma}_{k,m}} - 1} \frac{2M_1 - M_0}{\Gamma(1 - \hat{\gamma}_{k,m})}, \quad (5)$$

and the estimator $\hat{b}_{k,m}$ of b_m is

$$\hat{b}_{k,m} = M_0 + \hat{a}_{k,m} \frac{1 - \Gamma(1 - \hat{\gamma}_{k,m})}{\hat{\gamma}_{k,m}}, \quad (6)$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, $x > 0$ (Hosking, Wallis and Wood, 1985).

For $\hat{\gamma}_{k,m} = 0$ the estimators follow by continuity:

$$\hat{\gamma}_{k,m} = 0 \quad \text{if} \quad \frac{3M_2 - M_0}{2M_1 - M_0} = \frac{\log 3}{\log 2},$$

$$\hat{a}_{k,m} = \frac{2M_1 - M_0}{\log 2} \quad \text{and} \quad \hat{b}_{k,m} = M_0 + \hat{a}_{k,m} \Gamma'(1).$$

Note that $-\Gamma'(1)$ is Euler's constant.

Clearly, the given estimators are quite different from the ones in the POT case.

Remark 2.1. *There are other variants for M_r , e.g. $\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k+1}\right)^r X_{i,k}$, but they give theoretical problems.*

2.1. Asymptotic normality

Next we state conditions for the asymptotic normality of the mentioned estimators.

Theorem 2.1. *Assume that F is in the domain of attraction of an extreme value distribution G_γ with $\gamma < 1/2$ and that Condition 2.1 holds. Let $m = m_n \rightarrow \infty$ and $k = k_n \rightarrow \infty$ as $n \rightarrow \infty$, in such a way that $\sqrt{k}A(m) \rightarrow \lambda \in \mathbb{R}$. Then*

$$\sqrt{k} \left(\frac{(r+1)M_r - b_m}{a_m} - D_r(\gamma) \right) \rightarrow^d (r+1) \int_0^1 s^{r-1} (-\log s)^{-1-\gamma} B(s) ds + \lambda I_r(\gamma, \rho) =: Q_r, \quad (7)$$

as $n \rightarrow \infty$, jointly for $r = 0, 1, 2$, where \rightarrow^d means convergence in distribution, B is Brownian bridge,

$$D_r(\xi) = \frac{(r+1)\xi\Gamma(1-\xi) - 1}{\xi}, \quad \xi < 1$$

($D_r(0) = \log(r+1) - \Gamma'(1)$ as defined by continuity), and

$$I_r(\gamma, \rho) = \begin{cases} \frac{1}{\rho} (D_r(\gamma + \rho) - D_r(\gamma)), & \rho \neq 0, \\ D'_r(\gamma) = \frac{(r+1)^\gamma}{\gamma} \left(-\Gamma'(1-\gamma) + \log(r+1)\Gamma(1-\gamma) - \frac{D_r(\gamma)}{(r+1)^\gamma} \right), & \gamma \neq 0, \rho = 0, \\ D'_r(0) = \frac{1}{2} (\log^3(r+1) + \Gamma''(1) - 3\log(r+1)\Gamma'(1)), & \gamma = 0, \rho = 0. \end{cases}$$

Note that $\Gamma'(1-\gamma) = \int_0^\infty u^{-\gamma} e^{-u} (\log u) du$.

Remark 2.2. *Explicit (complicated) expressions for the limiting covariance matrix can be found in Hosking, Wallis and Wood (1985), cf. $v_{r,r}$, $v_{r,r+1}$ and $v_{r,r+s}$ (C.9)–(C.11) in their Appendix C. From there, $\text{var}(Q_r) = (r+1)^2 v_{rr}$ and $\text{cov}(Q_r, Q_s) = (r+1)(s+1)v_{r,s}$.*

Remark 2.3. *The condition $\sqrt{k}A(m) \rightarrow \lambda \in \mathbb{R}$ means that the growth of k_n , the number of blocks, is restricted with respect to the growth m_n , the size of a block, as $n \rightarrow \infty$. In particular this condition implies that $(\log k)/m \rightarrow 0$, as $n \rightarrow \infty$.*

The asymptotic normality of $\hat{\gamma}_{k,m}$, $\hat{a}_{k,m}$ and $\hat{b}_{k,m}$ follows from Theorem 2.1:

Theorem 2.2. *Under the conditions of Theorem 2.1, as $n \rightarrow \infty$,*

$$\sqrt{k}(\hat{\gamma}_{k,m} - \gamma) \rightarrow_d \frac{1}{\Gamma(1-\gamma)} \left(\frac{\log 3}{1-3^{-\gamma}} - \frac{\log 2}{1-2^{-\gamma}} \right)^{-1} \left\{ \frac{\gamma}{3^\gamma - 1} (Q_2 - Q_0) - \frac{\gamma}{2^\gamma - 1} (Q_1 - Q_0) \right\} =: \Delta,$$

$$\sqrt{k} \left(\frac{\hat{a}_{k,m}}{a_m} - 1 \right) \rightarrow_d \frac{\gamma}{(2^\gamma - 1)\Gamma(1-\gamma)} (Q_1 - Q_0) + \Delta \left\{ \frac{\log 2}{\gamma} \left(\frac{-\gamma}{1-2^{-\gamma}} + \frac{1}{\log 2} \right) + \frac{\Gamma'(1-\gamma)}{\Gamma(1-\gamma)} \right\} =: A$$

$$\sqrt{k} \frac{\hat{b}_{k,m} - b_m}{a_m} \rightarrow_d Q_0 + \frac{\gamma\Gamma'(1-\gamma) - 1 + \Gamma(1-\gamma)}{\gamma^2} \Delta + \frac{1 - \Gamma(1-\gamma)}{\gamma} A =: B;$$

where for $\gamma = 0$ the formulas should read as (defined by continuity):

$$\sqrt{k}\hat{\gamma}_{k,m} \rightarrow_d \left(\frac{\log 3}{2} - \frac{\log 2}{2} \right)^{-1} \left(\frac{1}{\log 3} (Q_2 - Q_0) - \frac{1}{\log 2} (Q_1 - Q_0) \right)$$

$$\begin{aligned} \sqrt{k} \left(\frac{\hat{a}_{k,m}}{a_m} - 1 \right) &\rightarrow_d \frac{1}{\log 2} (Q_1 - Q_0) + \Delta \left(\frac{\log 2}{2} + \Gamma'(1) \right) \\ \sqrt{k} \frac{\hat{b}_{k,m} - b_m}{a_m} &\rightarrow_d Q_0 - \Gamma''(1)\Delta + \Gamma'(1)\Lambda \end{aligned}$$

Remark 2.4. *The second order condition for the asymptotic normality in the POT case (see e.g. de Haan and Ferreira, Theorem 3.6.1), is formulated in terms of the function $U = (1/(1-F))^{\leftarrow}$, not the function V . For a comparison between the two second order conditions see Drees, de Haan and Li (2003).*

2.2. High quantile estimation

High quantile estimation is discussed in the next theorem. Let the total number of observations be n , that is, $n = mk$ as before. Let $m = m(n) \rightarrow \infty$, $k = k(n) \rightarrow \infty$ as $n \rightarrow \infty$. We want to estimate $x_n := F^{\leftarrow}(1 - p_n)$ with p_n small and write $x_n = V(1/(-\log(1 - p_n)))$.

Our estimator for x_n is

$$\hat{x}_{k,m} := \hat{b}_{k,m} + \hat{a}_{k,m} \frac{c_n^{\hat{\gamma}_{k,m}} - 1}{\hat{\gamma}_{k,m}}$$

with $c_n = (-m \log(1 - p_n))^{-1}$ and we obtain the following result:

Theorem 2.3. *Assume the conditions of Theorem 2.1 with the second order parameter ρ negative, or zero with γ negative. Moreover assume*

$$\lim_{n \rightarrow \infty} c_n = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\log c_n}{\sqrt{k}} = 0.$$

Then, as $n \rightarrow \infty$,

$$\frac{\sqrt{k}(\hat{x}_{k,m} - x_n)}{a_m q_\gamma(c_n)} \rightarrow_d \Delta + (\gamma_-)^2 B - \gamma_- \Lambda - \lambda \frac{\gamma_-}{\gamma_- + \rho}$$

where $\gamma_- := \min(0, \gamma)$ and

$$q_\gamma(t) := \int_1^t s^{\gamma-1} \log s \, ds.$$

Remark 2.5. *This is the result for a high quantile of the original distribution F . One may also want to estimate a high quantile of the distribution of the block maximum. In that case we need to estimate $x_n := V(m/(-\log(1 - p_n)))$. The result is as above with c_n replaced by mc_n .*

3. Proofs

Throughout this section Z represents a unit Fréchet random variable, i.e. one with distribution function $F(x) = e^{-1/x}$, $x > 0$, and $\{Z_{i,k}\}_{i=1}^k$ are the order statistics from the associated i.i.d. sample of size k , Z_1, \dots, Z_k . Similarly, $\{X_{i,k}\}_{i=1}^k$ represents the order statistics of the block maxima X_1, \dots, X_k from (1) and, $X_{\lceil u \rceil, k} := X_{r,k}$ for $r-1 < u \leq r$, $r = 1, \dots, k$. Recall the function V from Section 2. The following representation will be useful,

$$X =^d V(mZ). \quad (8)$$

We start by formulating a number of auxiliary results.

Lemma 3.1. 1. As $k \rightarrow \infty$,

$$(\log k)Z_{1,k} \rightarrow^P 1.$$

2. (Csörgő and Horváth 1993, p. 381) Let $0 < \nu < 1/2$. With $\{B_k\}_{k \geq 1}$ an appropriate sequence of Brownian bridges,

$$\sup_{1/(k+1) \leq s \leq k/(k+1)} \frac{s(-\log s)}{(s(1-s))^\nu} \left| \sqrt{k} \left((-\log s)Z_{\lceil ks \rceil, k} - 1 \right) - \frac{B_k(s)}{s(-\log s)} \right| = o_P(1),$$

as $k \rightarrow \infty$ ($\lceil u \rceil$ represents the smallest integer larger or equal to u).

3. Similarly, with $0 < \nu < 1/2$ for an appropriate sequence $\{B_k\}_{k \geq 1}$ of Brownian bridges and $\xi \in \mathbb{R}$,

$$\sup_{1/(k+1) \leq s \leq k/(k+1)} (s(1-s))^{-\nu} \left| \sqrt{k}s(-\log s)^{1+\xi} \left(\frac{Z_{\lceil ks \rceil, k}^\xi - 1}{\xi} - \frac{(-\log s)^{-\xi} - 1}{\xi} \right) - B_k(s) \right| = o_P(1),$$

as $k \rightarrow \infty$.

Lemma 3.2. Under Condition 2.1, there are functions $A(t) \sim A_0(t)$ and $a(t) = a_0(t)(1 + o(A_0(t)))$, as $t \rightarrow \infty$, such that for all $\varepsilon, \delta > 0$ there exists $t_0 = t_0(\varepsilon, \delta)$ such that for $t, tx > t_0$,

$$\left| \frac{\frac{V(tx) - V(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} - H_{\gamma, \rho}(x) \right| \leq \varepsilon \max(x^{\gamma+\rho+\delta}, x^{\gamma+\rho-\delta}). \quad (9)$$

Moreover,

$$\left| \frac{\frac{a(tx)}{a(t)} - x^\gamma}{A(t)} - x^\gamma \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon \max(x^{\gamma+\rho+\delta}, x^{\gamma+\rho-\delta}) \quad (10)$$

and

$$\left| \frac{A(tx)}{A(t)} - x^\rho \right| \leq \varepsilon \max(x^{\rho+\delta}, x^{\rho-\delta}).$$

Remark 3.1. This is an easily obtained variant of Theorem B.3.10 of de Haan and Ferreira (2006).

Note that,

$$H_{\gamma,\rho}(x) = \begin{cases} \frac{1}{\rho} \left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma} \right), & \rho \neq 0 \neq \gamma \\ \frac{1}{\gamma} \left(x^\gamma \log x - \frac{x^\gamma-1}{\gamma} \right), & \rho = 0 \neq \gamma \\ \frac{1}{\rho} \left(\frac{x^\rho-1}{\rho} - \log x \right), & \rho \neq 0 = \gamma \\ \frac{1}{2}(\log x)^2, & \rho = 0 = \gamma. \end{cases}$$

Proposition 3.1. Assume the conditions of Theorem 2.1. Let $0 < \varepsilon < 1/2$ and $\{X_{i,k}\}_{i=1}^k$ be the order statistics of the block maxima X_1, X_2, \dots, X_k . Then,

$$\begin{aligned} \sqrt{k} \left(\frac{X_{\lceil ks \rceil, k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) &= \frac{B_k(s)}{s(-\log s)^{1+\gamma}} + \sqrt{k} A(m) H_{\gamma,\rho} \left(\frac{1}{-\log s} \right) \\ &+ \left(s^{-1/2-\varepsilon} (1-s)^{-1/2-\gamma-\rho-\varepsilon} + B(s) s^{-1} (-\log s)^{-1-\gamma-\rho-\varepsilon} \right) o_P(1), \end{aligned}$$

as $n \rightarrow \infty$, where the $o_P(1)$ term is uniform for $1/(k+1) \leq s \leq k/(k+1)$.

Remark 3.2. This proposition should also be useful when analysing other estimators for the block maxima approach, like the maximum likelihood estimators.

Proof of Proposition 3.1. By representation (8),

$$\begin{aligned} &\frac{X_{\lceil ks \rceil, k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \\ &=_d \left(\frac{V(mZ_{\lceil ks \rceil, k}) - b_m}{a_m} - \frac{V\left(\frac{m}{-\log s}\right) - b_m}{a_m} \right) \\ &\quad + \left(\frac{V\left(\frac{m}{-\log s}\right) - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) \\ &= \text{I (random part)} + \text{II (bias part)}. \end{aligned}$$

We start with part I,

$$\begin{aligned} \text{I} &= \left\{ (-\log s)^{-\gamma} \frac{V\left((-\log s)Z_{\lceil ks \rceil, k} \frac{m}{-\log s}\right) - V\left(\frac{m}{-\log s}\right)}{a\left(\frac{m}{-\log s}\right)} \right\} \\ &\quad \times \left\{ \frac{a\left(\frac{m}{-\log s}\right)}{a(m)} (-\log s)^\gamma \right\} \\ &= \text{I.1} \times \text{I.2}. \end{aligned}$$

According to (10) of Lemma 3.2, for each $\varepsilon, \delta > 0$ there exists t_0 such that the factor I.2 is bounded (above and below) by

$$1 + A(m) \left\{ \frac{(-\log s)^{-\rho} - 1}{\rho} \pm \varepsilon \max \left((-\log s)^{-\rho+\delta}, (-\log s)^{-\rho-\delta} \right) \right\}$$

provided $m \geq t_0$ and $s \geq e^{-m/t_0}$. According to (9) of Lemma 3.2, for factor I.1 we have the bounds

$$\begin{aligned} & (-\log s)^{-\gamma} \frac{((-\log s)Z_{\lceil ks \rceil, k})^\gamma - 1}{\gamma} \\ & + A \left(\frac{m}{-\log s} \right) (-\log s)^{-\gamma} \times \{ H_{\gamma, \rho} ((-\log s)Z_{\lceil ks \rceil, k}) \\ & \pm \varepsilon \max \left(((-\log s)Z_{\lceil ks \rceil, k})^{\gamma+\rho+\delta}, ((-\log s)Z_{\lceil ks \rceil, k})^{\gamma+\rho-\delta} \right) \} \\ & = \text{I.1a} + \text{I.1b} \end{aligned}$$

provided $s \geq e^{-m/t_0}$ and $m/\log k \geq t_0$ (the latter inequality eventually holds true since $\sqrt{k}A(m)$ is bounded). Note that $m/\log k \geq t_0$ implies $mZ_{1,k} \geq 2t_0$ which implies (Lemma 3.1) $mZ_{\lceil ks \rceil, k} \geq 2t_0$ for all s .

For term I.1a we use Lemma 3.1.3:

$$\frac{(Z_{\lceil ks \rceil, k})^\gamma - 1}{\gamma} - \frac{(-\log s)^{-\gamma} - 1}{\gamma}$$

is bounded (above and below) by

$$\frac{1}{\sqrt{k}} \frac{B_k(s)}{s(-\log s)^{1+\gamma}} \pm \frac{\varepsilon}{\sqrt{k}} \frac{(s(1-s))^\nu}{s(-\log s)^{1+\gamma}},$$

for some $\varepsilon > 0$, $0 < \nu < 1/2$ and all $s \in [1/(k+1), k/(k+1)]$.

Next we turn to term I.1b. By Lemma 3.2, $(-\log s)^{-\gamma} A \left(\frac{m}{-\log s} \right)$ is bounded (above and below) by

$$A(m) \left\{ (-\log s)^{-\gamma-\rho} \pm \varepsilon \max \left((-\log s)^{-\gamma-\rho+\delta}, (-\log s)^{-\gamma-\rho-\delta} \right) \right\}$$

provided $s > e^{-m/t_0}$ and $m/\log k > t_0$. Furthermore for $\rho \neq 0 \neq \gamma$ and $s \in [1/(k+1), k/(k+1)]$, by Lemma 3.1.3,

$$\begin{aligned} & H_{\gamma, \rho} ((-\log s)Z_{\lceil ks \rceil, k}) \\ & = \frac{1}{\rho} \left\{ \frac{((-\log s)Z_{\lceil ks \rceil, k})^{\gamma+\rho} - 1}{\gamma + \rho} - \frac{((-\log s)Z_{\lceil ks \rceil, k})^\gamma - 1}{\gamma} \right\} \\ & = \frac{1}{\rho} \left\{ (-\log s)^{\gamma+\rho} \left[\frac{Z_{\lceil ks \rceil, k}^{\gamma+\rho} - 1}{\gamma + \rho} - \frac{(-\log s)^{-\gamma-\rho} - 1}{\gamma + \rho} \right] \right. \\ & \quad \left. - (-\log s)^\gamma \left[\frac{Z_{\lceil ks \rceil, k}^\gamma - 1}{\gamma} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right] \right\} \end{aligned}$$

is bounded by

$$\begin{aligned} & \frac{1}{\rho} \left\{ (-\log s)^{\gamma+\rho} \left[\frac{1}{\sqrt{k}} \frac{B_k(s)}{s(-\log s)^{1+\gamma+\rho}} \pm \frac{\varepsilon}{\sqrt{k}} \frac{(s(1-s))^\nu}{s(-\log s)^{1+\gamma+\rho}} \right] \right. \\ & \quad \left. - (-\log s)^\gamma \left[\frac{1}{\sqrt{k}} \frac{B_k(s)}{s(-\log s)^{1+\gamma}} \mp \frac{\varepsilon}{\sqrt{k}} \frac{(s(1-s))^\nu}{s(-\log s)^{1+\gamma}} \right] \right\} \\ & = \pm \frac{2\varepsilon}{\rho\sqrt{k}} \frac{(s(1-s))^\nu}{s(-\log s)}, \end{aligned}$$

and similarly for cases other than $\rho \neq 0 \neq \gamma$. The remaining part of I.1b, namely $\pm \varepsilon \max \left(((-\log s)Z_{\lceil ks \rceil, k})^{\gamma+\rho+\delta}, ((-\log s)Z_{\lceil ks \rceil, k})^{\gamma+\rho-\delta} \right)$, is similar.

Part II, by the inequalities of Lemma 3.2, is bounded by

$$A(m) \left\{ H_{\gamma, \rho} \left(\frac{1}{-\log s} \right) \pm \varepsilon \max \left((-\log s)^{-\gamma-\rho+\delta}, (-\log s)^{-\gamma-\rho-\delta} \right) \right\}$$

hence it contributes $\sqrt{k}A(m)H_{\gamma, \rho} \left(\frac{1}{-\log s} \right)$ to the result.

Collecting all the terms, one finds the result. \square

Proof of Theorem 2.1. Let, for $r = 0, 1, 2$,

$$J_k^{(r)}(s) = \frac{(\lceil ks \rceil - 1) \dots (\lceil ks \rceil - r)}{(k-1) \dots (k-r)}, \quad s \in [0, 1].$$

Note that $J_k^{(r)}(s) \rightarrow s^r$, as $k \rightarrow \infty$, uniformly in $s \in [0, 1]$, and,

$$\frac{1}{k} \sum_{i=1}^k \frac{(i-1) \dots (i-r)}{(k-1) \dots (k-r)} = \int_0^1 J_k^{(r)}(s) ds = \frac{1}{r+1} = \int_0^1 s^r ds.$$

Then,

$$\begin{aligned}
& \sqrt{k} \left(\frac{(r+1)M_r - b_m}{a_m} - \frac{(r+1)^\gamma \Gamma(1-\gamma) - 1}{\gamma} \right) \\
&= \sqrt{k} \left(\frac{(r+1) \int_0^1 X_{\lceil ks \rceil, k} J_k^{(r)}(s) ds - b_m}{a_m} - (r+1) \int_0^1 \frac{(-\log s)^{-\gamma} - 1}{\gamma} s^r ds \right) \\
&= \sqrt{k}(r+1) \int_0^1 \left(\frac{X_{\lceil ks \rceil, k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) J_k^{(r)}(s) ds \\
&\quad - \sqrt{k}(r+1) \int_0^1 \frac{(-\log s)^{-\gamma} - 1}{\gamma} (s^r - J_k^{(r)}(s)) ds \\
&= \sqrt{k}(r+1) \int_0^{1/(k+1)} \left(\frac{X_{\lceil ks \rceil, k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) J_k^{(r)}(s) ds \\
&\quad + \sqrt{k}(r+1) \int_{1/(k+1)}^{k/(k+1)} \left(\frac{X_{\lceil ks \rceil, k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) J_k^{(r)}(s) ds \\
&\quad + \sqrt{k}(r+1) \int_{k/(k+1)}^1 \left(\frac{X_{\lceil ks \rceil, k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) J_k^{(r)}(s) ds \\
&\quad - \sqrt{k}(r+1) \int_0^1 \frac{(-\log s)^{-\gamma} - 1}{\gamma} (s^r - J_k^{(r)}(s)) ds \\
&= \text{I1} + \text{I2} + \text{I3} + \text{I4}.
\end{aligned}$$

For I4: since $(s^r - J_k^{(r)}(s)) = O(1/k)$ uniformly in s , I4 = $O(1/\sqrt{k})$.

For I1, note that

$$\int_0^{1/(k+1)} \sqrt{k} \frac{X_{\lceil ks \rceil, k} - b_m}{a_m} s^r ds = o_P(1). \quad (11)$$

This follows since, the left-hand side of (11) equals, in distribution,

$$\frac{\sqrt{k}}{(k+1)^{r+1}} \frac{V(mZ_{1,k}) - V(m)}{a_m}$$

which, by Lemma 3.1.1, Lemma 3.2 and the fact that $m/\log k \rightarrow \infty$, is bounded (below and above) by

$$\frac{\sqrt{k}}{(k+1)^{r+1}} \left\{ \frac{Z_{1,k}^\gamma - 1}{\gamma} + A(m) H_{\gamma, \rho}(Z_{1,k}) \pm \varepsilon A(m) \max \left(Z_{1,k}^{\gamma+\rho+\delta}, Z_{1,k}^{\gamma+\rho-\delta} \right) \right\}.$$

This is easily seen to converge to zero in probability, since $Z_{1,k}^\xi/\sqrt{k} = \{(\log k) Z_{1,k}\}^\xi \log^{-\xi} k/\sqrt{k} \rightarrow^P 0$ for all real ξ and $\sqrt{k}A(m) \rightarrow \lambda$. Hence, I1 = $o_P(1)$.

Next we show that

$$\int_{k/(k+1)}^1 \sqrt{k} \frac{X_{\lceil ks \rceil, k} - b_m}{a_m} J_k^{(r)}(s) ds = o_P(1). \quad (12)$$

The left-hand side equals, in distribution (since $J_k^{(r)}(s) \equiv 1$ for $s \in (k(k+1)^{-1}, 1)$)

$$\left(1 - \frac{k}{k+1}\right) \sqrt{k} \frac{V(mZ_{k,k}) - V(m)}{a_m}.$$

Lemma 3.1 yields

$$\frac{V(mZ_{k,k}) - V(m)}{a_m} = \frac{Z_{k,k}^\gamma - 1}{\gamma} + A(m) \left\{ H_{\gamma,\rho}(Z_{k,k}) \pm \varepsilon Z_{k,k}^{\gamma+\rho+\delta} \right\}$$

which is (since $Z_{k,k}^\gamma/k^\gamma$ converges to a positive random variable) of the order $O_P(k^\gamma)$. Hence the integral is of order $(k+1)^{-1}\sqrt{k}k^\gamma$ which tends to zero since $\gamma < 1/2$.

Finally, I2 has the same asymptotic behaviour as

$$(r+1) \int_{1/(k+1)}^{k/(k+1)} \sqrt{k} \left(\frac{X_{[ks],k} - b_m}{a_m} - \frac{(-\log s)^{-\gamma} - 1}{\gamma} \right) s^r ds$$

which, by Proposition 3.1 tends to

$$(r+1) \int_0^1 s^{r-1} (-\log s)^{-1-\gamma} B(s) ds + \lambda(r+1) \int_0^1 H_{\gamma,\rho} \left(\frac{1}{-\log s} \right) s^r ds.$$

For the evaluation of the latter integral note that for $\xi < 1$,

$$(r+1) \int_0^1 s^r (-\log s)^{-\xi} ds = (r+1)^{\xi-1} \int_0^\infty v^{-\xi} e^{-v} dv = (r+1)^{\xi-1} \Gamma(1-\xi).$$

Moreover, note that

$$(r+1) \int_0^1 s^r \frac{(-\log s)^{-\xi} - 1}{\xi} ds = \frac{(r+1)^\xi \Gamma(1-\xi) - 1}{\xi}, \quad \xi < 1$$

($D_r(0) = \log(r+1) - \Gamma'(1)$ as defined by continuity), and

$$(r+1) \int_0^1 H_{\gamma,\rho} \left(\frac{1}{-\log s} \right) s^r ds = \begin{cases} \frac{1}{\rho} (D_r(\gamma + \rho) - D_r(\gamma)), & \rho \neq 0, \\ \frac{D_r'(\gamma)}{\gamma} = \frac{(r+1)^\gamma}{\gamma} \left(-\Gamma'(1-\gamma) + \log(r+1)\Gamma(1-\gamma) - \frac{D_r(\gamma)}{(r+1)^\gamma} \right), & \gamma \neq 0, \rho = 0, \\ \frac{D_r'(0)}{\frac{1}{2}} = \frac{1}{2} (\log^3(r+1) + \Gamma''(1) - 3\log(r+1)\Gamma'(1)), & \gamma = 0, \rho = 0. \end{cases}$$

□

Proof of Theorem 2.2. From Theorem 2.1 we obtain,

$$\begin{aligned}\sqrt{k} \left(\frac{2M_1 - M_0}{a_m} - \frac{2^\gamma - 1}{\gamma} \Gamma(1 - \gamma) \right) &\rightarrow^d Q_1 - Q_0 \\ \sqrt{k} \left(\frac{3M_2 - M_0}{a_m} - \frac{3^\gamma - 1}{\gamma} \Gamma(1 - \gamma) \right) &\rightarrow^d Q_2 - Q_0\end{aligned}$$

hence, by Cramér's delta method,

$$\begin{aligned}\sqrt{k} \left(\frac{3^{\hat{\gamma}_{k,m}} - 1}{2^{\hat{\gamma}_{k,m}} - 1} - \frac{3^\gamma - 1}{2^\gamma - 1} \right) &= \sqrt{k} \left(\frac{3M_2 - M_0}{2M_1 - M_0} - \frac{3^\gamma - 1}{2^\gamma - 1} \right) \\ &\rightarrow^d \frac{1}{\Gamma(1 - \gamma)} \frac{3^\gamma - 1}{2^\gamma - 1} \left(\frac{\gamma}{3^\gamma - 1} (Q_2 - Q_0) - \frac{\gamma}{2^\gamma - 1} (Q_1 - Q_0) \right).\end{aligned}$$

It follows that $\hat{\gamma}_{k,m} \xrightarrow{P} \gamma$ and hence

$$\sqrt{k} \left(\frac{r^{\hat{\gamma}_{k,m}} - 1}{r^\gamma - 1} - 1 \right) = \sqrt{k} \frac{r^{\hat{\gamma}_{k,m} - \gamma} - 1}{1 - r^{-\gamma}}$$

has the same limit distribution as

$$\sqrt{k} (\hat{\gamma}_{k,m} - \gamma) \frac{\log r}{1 - r^{-\gamma}}, \quad r = 2, 3.$$

It follows that

$$\begin{aligned}\sqrt{k} \left(\frac{3^{\hat{\gamma}_{k,m}} - 1}{2^{\hat{\gamma}_{k,m}} - 1} - \frac{3^\gamma - 1}{2^\gamma - 1} \right) \\ = \frac{3^\gamma - 1}{2^\gamma - 1} \left[\sqrt{k} \left(\frac{3^{\hat{\gamma}_{k,m}} - 1}{3^\gamma - 1} - 1 \right) - \sqrt{k} \left(\frac{2^{\hat{\gamma}_{k,m}} - 1}{2^\gamma - 1} - 1 \right) \right]\end{aligned}$$

has the same limit distribution as

$$\frac{3^\gamma - 1}{2^\gamma - 1} \sqrt{k} (\hat{\gamma}_{k,m} - \gamma) \left(\frac{\log 3}{1 - 3^{-\gamma}} - \frac{\log 2}{1 - 2^{-\gamma}} \right)$$

and, consequently,

$$\begin{aligned}\sqrt{k} (\hat{\gamma}_{k,m} - \gamma) \\ \rightarrow^d \frac{1}{\Gamma(1 - \gamma)} \left(\frac{\log 3}{1 - 3^{-\gamma}} - \frac{\log 2}{1 - 2^{-\gamma}} \right)^{-1} \left(\frac{\gamma}{3^\gamma - 1} (Q_2 - Q_0) - \frac{\gamma}{2^\gamma - 1} (Q_1 - Q_0) \right).\end{aligned}$$

For the asymptotic distribution of $\hat{a}_{k,m}$ we write,

$$\begin{aligned}\sqrt{k} \left(\frac{\hat{a}_{k,m}}{a_m} - 1 \right) &= \frac{\hat{\gamma}_{k,m}}{(2^{\hat{\gamma}_{k,m}} - 1) \Gamma(1 - \hat{\gamma}_{k,m})} \\ \left\{ \sqrt{k} \left(\frac{2M_1 - M_0}{a_m} - \frac{2^\gamma - 1}{\gamma} \Gamma(1 - \gamma) \right) + \sqrt{k} \left(\frac{2^\gamma - 1}{\gamma} \Gamma(1 - \gamma) - \frac{2^{\hat{\gamma}_{k,m}} - 1}{\hat{\gamma}_{k,m}} \Gamma(1 - \hat{\gamma}_{k,m}) \right) \right\},\end{aligned}$$

and the statement follows e.g. by Cramér's delta method.

For the asymptotic distribution of $\hat{b}_{k,m}$ we write,

$$\begin{aligned} \sqrt{k} \left(\frac{\hat{b}_{k,m} - b_m}{a_m} \right) &= \sqrt{k} \left(\frac{M_0 - b_m}{a_m} - \frac{\Gamma(1-\gamma) - 1}{\gamma} \right) \\ &- \frac{\hat{a}_{k,m}}{a_m} \sqrt{k} \left(\frac{\Gamma(1-\hat{\gamma}_{k,m}) - 1}{\hat{\gamma}_{k,m}} - \frac{\Gamma(1-\gamma) - 1}{\gamma} \right) + \frac{\Gamma(1-\gamma) - 1}{\gamma} \sqrt{k} \left(\frac{\hat{a}_{k,m}}{a_m} - 1 \right) \end{aligned}$$

and the statement follows e.g. by Cramér's delta method. \square

Proof of Theorem 2.3. The proof follows the line of the comparable result for the peaks-over-threshold method (see e.g. de Haan and Ferreira 2006, Chapter 4.3)

$$\begin{aligned} \frac{\sqrt{k}(\hat{x}_{k,m} - x_n)}{a_m q_\gamma(c_n)} &= \frac{\sqrt{k}}{a_m q_\gamma(c_n)} \left(\hat{b}_{k,m} + \hat{a}_{k,m} \frac{c_n^{\hat{\gamma}_{k,m}} - 1}{\hat{\gamma}_{k,m}} - V \left(\frac{1}{-\log(1-p_n)} \right) \right) \\ &= \frac{\sqrt{k}}{q_\gamma(c_n)} \frac{\hat{b}_{k,m} - b_m}{a_m} + \frac{\hat{a}_{k,m}}{a_m} \frac{\sqrt{k}}{q_\gamma(c_n)} \left(\frac{c_n^{\hat{\gamma}_{k,m}} - 1}{\hat{\gamma}_{k,m}} - \frac{c_n^\gamma - 1}{\gamma} \right) \\ &\quad - \frac{\sqrt{k}}{q_\gamma(c_n)} \left(\frac{V \left(\frac{m}{-m \log(1-p_n)} \right) - V(m)}{a_m} - \frac{c_n^\gamma - 1}{\gamma} \right) + \frac{c_n^\gamma - 1}{\gamma q_\gamma(c_n)} \sqrt{k} \left(\frac{\hat{a}_{k,m}}{a_m} - 1 \right). \end{aligned}$$

Similarly as on pages 135–137 of de Haan and Ferreira (2006) this converges in distribution to

$$\Delta + (\gamma_-)^2 B - \gamma_- A - \lambda \frac{\gamma_-}{\gamma_- + \rho}.$$

\square

Acknowledgements. Research partially supported by FCT Project PTDC /MAT /112770 /2009 and FCT- PEst-OE/MAT/UI0006/2011.

We thank Holger Drees for a useful suggestion.

References

- [1] Caires, S. (2009) A comparative simulation study of the annual maxima and the peaks-over-threshold methods. SBW-Belastingen: subproject 'Statistics'. Deltares Report 1200264-002.
- [2] Csörgő, M. and Horváth, L. (1993) Weighted Approximations in Probability and Statistics. John Wiley & Sons, Chichester, England
- [3] Dombry, C. (2013) Maximum likelihood estimators for the extreme value index based on the block maxima method: arXiv:1301.5611
- [4] Drees, H. (1998) On smooth statistical tail functionals. Scand. J. Statist. **25**, 187–210.
- [5] Gumbel, E. (1958) Statistics of Extremes, Columbia University Press.

- [6] de Haan, L. and Ferreira, A. (2006) *Extreme Value Theory: An Introduction*. Springer, Boston.
- [7] Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985) Estimation of the Generalized Extreme-Value Distribution by the Method of Probability Weighted Moments. *Technometrics* **27**, 251–261.
- [8] Pickands, J. III (1975) Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119-131.
- [9] Prescott, P. and Walden, A.T. (1980) Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika* **67**, 723-724.