

Weak Convergence of the Empirical Mean Excess Process with Application to Estimate the Negative Tail Index

Jürg Hüsler · Deyuan Li

Received: 29 March 2007 / Revised: 23 November 2007 /
Accepted: 10 December 2007 / Published online: 28 February 2008
© Springer Science + Business Media, LLC 2007

Abstract Let Y_i , $1 \leq i \leq n$ be i.i.d. random variables with the generalized Pareto distribution $W_{\gamma, \sigma}$ with $\gamma < 0$. We define the empirical mean excess process with respect to $\{Y_i, 1 \leq i \leq n\}$ as in Eq. 2.1 (see below) and investigate its weak convergence. As an application, two new estimators of the negative tail index γ are constructed based on the linear regression to the empirical mean excess function and their consistency and asymptotic normality are obtained.

Keywords Mean excess function · Tail index · Linear regression · Empirical mean excess process · Goodness-of-fit test

AMS 2000 Subject Classification 62G32 · 60G70

1 Introduction

Suppose X, X_1, X_2, \dots, X_n are independently identically distributed (i.i.d.) random variables with distribution function F , which belongs to the max-domain of an extreme value distribution G_γ with $\gamma \in \mathbb{R}$, denoted by $F \in D(G_\gamma)$, i.e. there exist sequences $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$F^n(a_n x + b_n) \rightarrow G_\gamma(x) := \exp\left(-(1 + \gamma x)^{-1/\gamma}\right)$$

J. Hüsler (✉) · D. Li
Department of Mathematical Statistics, University of Bern, Bern, Switzerland
e-mail: juerg.huesler@stat.unibe.ch

D. Li
e-mail: deyuan.li@stat.unibe.ch

for $1 + \gamma x > 0$. Here, γ is the so-called extreme value index. $G_\gamma(x)$ is interpreted as $\exp(-e^{-x})$ for $\gamma = 0$. In case of $\gamma < 0$, $x_f := \sup\{x : F(x) < 1\} < \infty$. Then it follows that (e.g. see Resnick 1987)

$$\lim_{t \rightarrow 0} \frac{\bar{F}(x_f - tx)}{\bar{F}(x_f - t)} = x^{-1/\gamma}$$

for all $x > 0$, where $\bar{F} = 1 - F$. The excess Y_X , of u by X for large u , is

$$Y_X = Y_X(u) := X - u | X > u.$$

For $y > 0$ and $y + u \leq x_f$, it is known that for u close to x_f

$$\begin{aligned} P(Y_X \leq y) &= P(X - u \leq y | X > u) = \frac{\bar{F}(u) - \bar{F}(y + u)}{\bar{F}(u)} \\ &= 1 - \frac{\bar{F}(x_f - (x_f - y - u))}{\bar{F}(x_f - (x_f - u))} \approx 1 - \left(1 + \gamma \frac{y}{\sigma}\right)^{-1/\gamma} =: W_{\gamma,\sigma}(y) \end{aligned}$$

with $\sigma = \gamma(u - x_f) > 0$. Thus the distribution of Y_X converges to the generalized Pareto distribution $W_{\gamma,\sigma}$.

Let $u = X_{n-k,n}$, where $k = k(n) \in \mathbb{N}$ such that $k \rightarrow \infty$, $k/n \rightarrow 0$ as $n \rightarrow \infty$, and define

$$\tilde{Y}_{i,k} = X_{n-k+i,n} - X_{n-k,n}, \quad i = 1, \dots, k. \tag{1.1}$$

Then $\{\tilde{Y}_{i,k} : i = 1, 2, \dots, k\}$ can be roughly regarded as a sample from the distribution $W_{\gamma,\sigma}$.

Assume $Y \sim W_{\gamma,\sigma}$. The mean excess function of Y is given by $e_Y(t) := E(Y - t | Y > t)$ for $0 < t < -\sigma/\gamma$. Straightforward calculation shows that

$$e_Y(t) = \frac{\sigma}{1 - \gamma} + \frac{\gamma}{1 - \gamma} t =: \beta_0 + \beta_1 t. \tag{1.2}$$

Relation (1.2) can be used to estimate β_0 and β_1 by linear regression on the empirical mean excess function, and thus to obtain estimators for γ and σ . This idea was presented also in Beirlant et al. (1996) and Reiss and Thomas (2007). More precisely, assume Y_1, Y_2, \dots, Y_k are i.i.d. random variables distributed from the distribution $W_{\gamma,\sigma}$. Then the empirical mean excess function is defined by

$$\hat{e}_Y(t) = \frac{1}{\#\{i : Y_i > t, 1 \leq i \leq k\}} \sum_{i=1}^k (Y_i - t) I_{\{Y_i > t\}}.$$

Denoting the order statistics of Y_i 's by $0 = Y_{0,k} \leq Y_{1,k} \leq Y_{2,k} \leq \dots \leq Y_{k,k}$ and putting $t = Y_{i,k}$, we have

$$Z_i := \hat{e}_Y(Y_{i,k}) = \frac{1}{k - i} \sum_{j=i+1}^k Y_{j,k} - Y_{i,k}, \quad i = 0, 1, \dots, k - 1. \tag{1.3}$$

It is easy to check that $E(Z_i|Y_{i,k}) = \beta_0 + \beta_1 Y_{i,k}$ for $i = 0, 1, \dots, k - 1$. Thus, all the points $(Y_{i,k}, Z_i), i = 0, 1, \dots, k - 1$, can be used to estimate β_0 and β_1 by least squares estimation as in linear regression models, i.e.

$$\hat{\beta}_1 = \frac{\sum_{i=0}^{k-1} (Y_{i,k} - \bar{Y})(Z_i - \bar{Z})}{\sum_{i=0}^{k-1} (Y_{i,k} - \bar{Y})^2} \quad \text{and} \quad \hat{\beta}_0 = \bar{Z} - \hat{\beta}_1 \bar{Y}, \tag{1.4}$$

where \bar{Y} and \bar{Z} denote the means of $\{Y_{0,k}, Y_{1,k}, \dots, Y_{k-1,k}\}$ and $\{Z_0, Z_1, \dots, Z_{k-1}\}$, respectively. Consequently, by Eq. 1.2, the least squares estimators for γ and σ are

$$\hat{\gamma} = \frac{\hat{\beta}_1}{1 + \hat{\beta}_1} \quad \text{and} \quad \hat{\sigma} = \frac{\hat{\beta}_0}{1 + \hat{\beta}_1}. \tag{1.5}$$

Recall that $\{\tilde{Y}_{i,k} : i = 1, 2, \dots, k\}$ can be roughly regarded as a sample from the distribution $W_{\gamma,\sigma}$. We define the mean excess estimator of the extreme index by

$$\hat{\gamma}^X = \frac{\hat{\beta}_1^X}{1 + \hat{\beta}_1^X}, \tag{1.6}$$

where

$$\hat{\beta}_1^X = \frac{\sum_{i=0}^{k-1} (\tilde{Y}_{i,k} - \bar{\tilde{Y}})(\tilde{Z}_i - \bar{\tilde{Z}})}{\sum_{i=0}^{k-1} (\tilde{Y}_{i,k} - \bar{\tilde{Y}})^2} \tag{1.7}$$

with notation $\tilde{Y}_{0,k} = 0, \tilde{Z}_i = \frac{1}{k-i} \sum_{j=i+1}^k \tilde{Y}_{j,k} - \tilde{Y}_{i,k}, \bar{\tilde{Y}} = k^{-1} \sum_{i=0}^{k-1} \tilde{Y}_{i,k}$ and $\bar{\tilde{Z}} = k^{-1} \sum_{i=0}^{k-1} \tilde{Z}_i$.

In this paper, we investigate the asymptotic normality of the mean excess estimator $\hat{\gamma}^X$. The rest of this paper is organized as follows. In Section 2, we present the main results. Extension of the mean excess estimator and comparison with other estimators are discussed in Section 3. The proofs of the main results are shown in Section 4.

2 Main Results

In order to investigate the asymptotic normality of the mean excess estimator $\hat{\gamma}^X$, we need some asymptotic results on $\hat{\beta}_1$ (defined in Eq. 1.4 and based on the sample from the generalized Pareto distribution $W_{\gamma,\sigma}$). The asymptotic behavior of $\hat{\beta}_1$ is based on a suitably defined empirical mean excess process T_k as follows.

Let $U \sim U_\alpha[0, 1]$ with $\alpha > 0$ if $P(U \leq x) = x^\alpha$ for $0 \leq x \leq 1$. In case of $\alpha = 1$, it is the uniform distribution on $[0, 1]$ usually denoted by $U[0, 1]$. Let U_1, U_2, \dots, U_k be i.i.d. random variables with distribution $U_\alpha[0, 1]$ for some $\alpha > 0$ with order statistics $0 = U_{0,k} \leq U_{1,k} \leq \dots \leq U_{k,k} \leq U_{k+1,k} = 1$.

For fixed $i : 2 \leq i < k$, by Rényi representation for the uniform order statistics, $\{U_{j,k}/U_{i,k} : 1 \leq j < i\}$ can be regarded as a sample of size $i - 1$ from $U_\alpha[0, 1]$ with mean $(1 - \gamma)^{-1}$ and variance

$$\sigma_\gamma^2 := \frac{\gamma^2}{(1 - 2\gamma)(1 - \gamma)^2},$$

where $\gamma = -1/\alpha < 0$. Thus, by central limit theorem, for $i \leq k$

$$\frac{1}{\sigma_\gamma \sqrt{i-1}} \sum_{j=1}^{i-1} \left(\frac{U_{j,k}}{U_{i,k}} - \frac{1}{1-\gamma} \right)$$

converges in distribution to a standard normal random variable as $i \rightarrow \infty$. More generally, we consider the process T_k defined by

$$T_k(t) := \sigma_\gamma^{-1} k^{-1/2} \sum_{j=1}^{[kt]-1} \left(\frac{U_{j,k}}{U_{[kt],k}} - \frac{1}{1-\gamma} \right), \quad t \in [0, 1]. \tag{2.1}$$

The process T_k is called the *empirical mean excess process*. In case of $\gamma = -1$, T_k is called the *uniform empirical mean excess process*. For fixed $t \in [0, 1]$, it follows by central limit theory that $T_k(t)$ converges in distribution to a normal random variable with mean zero and variance t . We need to investigate the weak convergence of the process T_k to derive the limiting behavior of the estimates and test statistics mentioned in the beginning.

To specify the relation between $\hat{\beta}_1$ and the process T_k , let $U_i = 1 + \gamma \sigma^{-1} Y_i$ for $i = 1, 2, \dots, k$. Hence U_1, U_2, \dots, U_k are i.i.d. random variables with distribution $U_\alpha[0, 1]$, where $\alpha = -1/\gamma > 0$, and $Y_{i,k} = \sigma(U_{k-i+1,k} - 1)/\gamma$ for $i = 0, 1, \dots, k$. Thus for $i = 0, 1, \dots, k - 1$,

$$Z_i = \frac{\sigma}{\gamma} \left(\frac{1}{k-i} \sum_{j=i+1}^k U_{k-j+1,k} - U_{k-i+1,k} \right) =: \frac{\sigma}{\gamma} V_{k-i+1},$$

and by Eq. 1.4,

$$\hat{\beta}_1 = \frac{\sum_{i=0}^{k-1} (U_{k-i+1,k} - \bar{U})(V_{k-i+1} - \bar{V})}{\sum_{i=0}^{k-1} (U_{k-i+1,k} - \bar{U})^2} = \frac{\sum_{i=2}^{k+1} (U_{i,k} - \bar{U})(V_i - \bar{V})}{\sum_{i=2}^{k+1} (U_{i,k} - \bar{U})^2} \tag{2.2}$$

and

$$\hat{\beta}_0 = \frac{\sigma}{\gamma} (\bar{V} - \hat{\beta}_1(\bar{U} - 1)), \tag{2.3}$$

where \bar{U} and \bar{V} denote the mean of $\{U_{2,k}, U_{3,k}, \dots, U_{k+1,k}\}$ and of $\{V_2, V_3, \dots, V_{k+1}\}$, respectively.

Consider the numerator of the right hand side of Eq. 2.2. Recall that for $i \leq k$ and as $i \rightarrow \infty$

$$\frac{V_i}{U_{i,k}} = \frac{1}{i-1} \sum_{j=1}^{i-1} \left(\frac{U_{j,k}}{U_{i,k}} - 1 \right) = \frac{1}{i-1} \sum_{j=1}^{i-1} \left(\frac{U_{j,k}}{U_{i,k}} - \frac{1}{1-\gamma} \right) + \beta_1 \xrightarrow{P} \beta_1.$$

Thus for large i , $V_i = \frac{\gamma}{1-\gamma} U_{i,k} + o_P(1)$, and hence as $k \rightarrow \infty$

$$\sum_{i=2}^{k+1} (U_{i,k} - \bar{U})(V_i - \bar{V}) = \beta_1 \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})^2 + o_P(k).$$

So, in order to derive the consistency and asymptotic normality of $\hat{\beta}_0$ and $\hat{\beta}_1$, we need to analyze the asymptotic behavior of the process T_k .

For the empirical excess process we have the following result, which is a key to obtain the asymptotic normality of $\hat{\gamma}^X$.

Theorem 2.1 *Suppose U_1, U_2, \dots, U_k are i.i.d. random variables with distribution $U_\alpha[0, 1]$ where $\alpha > 0$. Let $\gamma = -1/\alpha$ and the sequence of processes $\{T_k, k \geq 1\}$ be defined as in Eq. 2.1. Then*

$$\left\{ T_k(t), t \in [0, 1] \right\} \xrightarrow{d} \left\{ \sigma_\gamma^{-1} \left(\frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B(u) du - \frac{1}{1+\alpha} B(t) \right), t \in [0, 1] \right\},$$

where B is a standard Brownian motion.

We are interested in the limiting process (in distribution) of $\{T_k, k \geq 1\}$. Let

$$W(t) = \sigma_\gamma^{-1} \left(\frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B(u) du - \frac{1}{1+\alpha} B(t) \right), \quad t \in [0, 1]. \tag{2.4}$$

W is well defined in the neighborhood of zero and $W(0) := \lim_{t \downarrow 0} W(t) = 0$. Simple but tedious calculations imply that W is a Gaussian process satisfying $EW(t) = 0$ and $EW(t)W(s) = t(t/s)^{1/\alpha}$ for $0 \leq t \leq s \leq 1$. As $\alpha \rightarrow \infty$, W converges to a standard Brownian motion.

Now let

$$\tilde{W}(t) = -\frac{1}{\sigma_\gamma(1+\alpha)t^{1/\alpha}} \int_0^t u^{1/\alpha} dB(u), \quad t \in [0, 1].$$

It is easy to check that \tilde{W} is a Gaussian process and also satisfies $E\tilde{W}(t) = 0$ and $E\tilde{W}(t)\tilde{W}(s) = t(t/s)^{1/\alpha}$ for $0 \leq t \leq s \leq 1$. Thus $W \stackrel{d}{=} \tilde{W}$. On the other hand, by partial integration, \tilde{W} can be rewritten as

$$\tilde{W}(t) = \sigma_\gamma^{-1} \left(\frac{1}{\alpha(1+\alpha)t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B(u) du - \frac{1}{1+\alpha} B(t) \right). \tag{2.5}$$

Note that the denominators of the integral parts in Eqs. 2.4 and 2.5 are different! We have the following interesting finding as a byproduct.

Corollary 2.1 *Let B be a standard Brownian motion and $\alpha > 0$. Then*

$$\begin{aligned} & \left\{ \frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B(u) du - \frac{1}{1+\alpha} B(t), t \geq 0 \right\} \\ & \stackrel{d}{=} \left\{ \frac{1}{\alpha(1+\alpha)t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B(u) du - \frac{1}{1+\alpha} B(t), t \geq 0 \right\}. \end{aligned}$$

In particular, for $\alpha = 1$

$$\left\{ \frac{1}{t} \int_0^t B(u) du - \frac{1}{2} B(t), t \geq 0 \right\} \stackrel{d}{=} \left\{ \frac{1}{2t} \int_0^t B(u) du - \frac{1}{2} B(t), t \geq 0 \right\}.$$

Based on Theorem 2.1, we obtain the asymptotic normality of $\hat{\beta}_1$ and other estimators.

Theorem 2.2 Suppose that Y_1, Y_2, \dots, Y_k are i.i.d. random variables with the distribution $W_{\gamma,\sigma}$ where $\gamma < 0$ and $\sigma > 0$. Let $\hat{\beta}_1, \hat{\beta}_0, \hat{\gamma}$ and $\hat{\sigma}$ are defined as in Eqs. 1.4 and 1.5. Then

$$k^{1/2} \begin{pmatrix} \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_0 - \beta_0 \end{pmatrix} \xrightarrow{d} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} \tag{2.6}$$

and

$$k^{1/2} \begin{pmatrix} \hat{\gamma} - \gamma \\ \hat{\sigma} - \sigma \end{pmatrix} \xrightarrow{d} \begin{pmatrix} (1 - \gamma)^2 N_1 \\ (1 - \gamma)(N_2 - \sigma N_1) \end{pmatrix}, \tag{2.7}$$

as $k \rightarrow \infty$, where W is a Gaussian process on $[0, 1]$ such that $EW(t) = 0$ and $EW(t)W(s) = t(t/s)^{-\gamma}$ for any $0 \leq t \leq s \leq 1$, and

$$N_1 = \sigma_\gamma^{-1} \int_0^1 \left(x^{-\gamma} - \frac{1}{1 - \gamma} \right) x^{-\gamma-1} W(x) dx,$$

$$N_2 = \frac{\sigma}{\gamma} \left[\sigma_\gamma \int_0^1 x^{-\gamma-1} W(x) dx - \frac{\gamma}{1 - \gamma} \sigma_\gamma^{-1} \int_0^1 \left(x^{-\gamma} - \frac{1}{1 - \gamma} \right) x^{-\gamma-1} W(x) dx \right].$$

Now let us turn to the mean excess estimator $\hat{\gamma}^X$. Let $U(t) = F^{\leftarrow}(1 - 1/t)$ for $t \geq 1$. Then for $\gamma < 0$, $U(\infty) = x_F < \infty$ and

$$\lim_{t \rightarrow \infty} \frac{U(\infty) - U(tx)}{U(\infty) - U(t)} \rightarrow x^\gamma, \quad x > 0$$

(e.g. see Resnick 1987). In order to obtain the asymptotic normality of $\hat{\gamma}^X$, we further require that F satisfies the following second order condition:

$$\lim_{t \rightarrow \infty} \frac{\frac{U(\infty) - U(tx)}{U(\infty) - U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}, \quad x > 0, \tag{2.8}$$

where $\rho \leq 0$ and A is an eventually positive or negative function such that $|A| \in RV(\rho)$ and $\lim_{t \rightarrow \infty} A(t) = 0$. This condition is common in extreme value theory (e.g. see de Haan and Stadtmüller 1996).

Theorem 2.3 Suppose X_1, X_2, \dots, X_n are i.i.d. random variables with the distribution function F such that $F \in D(G_\gamma)$ with $\gamma < 0$. Let $\hat{\beta}_1^X$ and $\hat{\gamma}^X$ are defined as in Eqs. 1.7 and 1.6, respectively, and assume the conditions (2.8) and $k^{1/2} A(n/k) \rightarrow \lambda \in \mathbb{R}$. Then

$$k^{1/2} \left(\hat{\beta}_1^X - \beta_1 \right) \xrightarrow{d} \Lambda + N_1 \quad \text{and} \quad k^{1/2} \left(\hat{\gamma}^X - \gamma \right) \xrightarrow{d} \Lambda + (1 - \gamma)^2 N_1,$$

as $n \rightarrow \infty$, where N_1 is the same as in Theorem 2.2 and

$$\Lambda = \lambda \sigma_\gamma^{-2} \int_0^1 \left(x^{-\gamma} - \frac{1}{1 - \gamma} \right) x^{-\gamma-1} \left(\int_0^x \left(\frac{y}{x} \right)^\gamma \frac{y^\rho - x^\rho}{\rho} dy \right) dx.$$

For $\rho = 0$, $(y^\rho - x^\rho)/\rho$ is interpreted as $\log(y/x)$.

In Theorem 2.3, Λ determines the bias of the estimates. If k is selected such that $k^{1/2} A(n/k) \rightarrow 0$, then the bias vanishes asymptotically.

3 Extension and Comparison

3.1 Extension

In this subsection we extend the mean excess estimator of the negative tail index. First note that $\hat{\beta}_1$ defined in Eq. 1.4 can be written as

$$\hat{\beta}_1 = \frac{\sum_{i=0}^{k-1} Y_{i,k} Z_i k^{-1} - \left(\sum_{i=0}^{k-1} Y_{i,k} k^{-1}\right) \left(\sum_{i=0}^{k-1} Z_i k^{-1}\right)}{\sum_{i=0}^{k-1} Y_{i,k}^2 k^{-1} - \left(\sum_{i=0}^{k-1} Y_{i,k} k^{-1}\right)^2}.$$

Let $\phi(x)$ be a nonnegative function on $[0, 1]$ such that $\int_0^1 \phi(x)dx = 1$. Now we replace the equal weight k^{-1} in the formula above by $\phi(i/k)k^{-1}$ and obtain a weighted mean excess estimator of β_1 , namely $\hat{\beta}_1(\phi)$,

$$\hat{\beta}_1(\phi) = \frac{\frac{1}{k} \sum_{i=0}^{k-1} Y_{i,k} Z_i \phi(i/k) - \left(\frac{1}{k} \sum_{i=0}^{k-1} Y_{i,k} \phi(i/k)\right) \left(\frac{1}{k} \sum_{i=0}^{k-1} Z_i \phi(i/k)\right)}{\frac{1}{k} \sum_{i=0}^{k-1} Y_{i,k}^2 \phi(i/k) - \left(\frac{1}{k} \sum_{i=0}^{k-1} Y_{i,k} \phi(i/k)\right)^2}.$$

Consequently, the weighted mean excess estimator for the negative tail index γ is defined by

$$\hat{\gamma}(\phi) = \frac{\hat{\beta}_1(\phi)}{1 + \hat{\beta}_1(\phi)}.$$

Similar to above, we also define the weighted mean excess estimator $\hat{\gamma}^X(\phi)$ for the negative extreme value index, i.e.

$$\hat{\gamma}^X(\phi) = \frac{\hat{\beta}_1^X(\phi)}{1 + \hat{\beta}_1^X(\phi)},$$

where

$$\hat{\beta}_1^X(\phi) = \frac{\frac{1}{k} \sum_{i=0}^{k-1} \tilde{Y}_{i,k} \tilde{Z}_i \phi(i/k) - \left(\frac{1}{k} \sum_{i=0}^{k-1} \tilde{Y}_{i,k} \phi(i/k)\right) \left(\frac{1}{k} \sum_{i=0}^{k-1} \tilde{Z}_i \phi(i/k)\right)}{\frac{1}{k} \sum_{i=0}^{k-1} \tilde{Y}_{i,k}^2 \phi(i/k) - \left(\frac{1}{k} \sum_{i=0}^{k-1} \tilde{Y}_{i,k} \phi(i/k)\right)^2}.$$

The asymptotic normalities of $\hat{\beta}_1(\phi)$, $\hat{\gamma}(\phi)$, $\hat{\beta}_1^X(\phi)$ and $\hat{\gamma}^X(\phi)$ are shown in the following theorem without proof.

Theorem 3.1 *Let ϕ be nonnegative on $[0, 1]$ such that $\int_0^1 \phi(x)dx = 1$.*

- (1) *Assume the conditions in Theorem 2.2. Then $\sqrt{k}(\hat{\beta}_1(\phi) - \beta_1) \xrightarrow{d} N_1(\phi)$ and $\sqrt{k}(\hat{\gamma}(\phi) - \gamma) \xrightarrow{d} (1 - \gamma)^2 N_1(\phi)$, where*

$$N_1(\phi) = \frac{\sigma_\gamma \int_0^1 (x^{-\gamma} - \int_0^1 t^{-\gamma} \phi(1-t)dt) x^{-\gamma-1} \phi(1-x) W(x) dx}{\int_0^1 x^{-2\gamma} \phi(1-x) dx - \left(\int_0^1 x^{-\gamma} \phi(1-x) dx\right)^2}.$$

(2) Assume the conditions in Theorem 2.3. Then $\sqrt{k}(\hat{\beta}_1^X(\phi) - \beta_1) \xrightarrow{d} \Lambda(\phi) + N_1(\phi)$ and $\sqrt{k}(\hat{\gamma}^X(\phi) - \gamma) \xrightarrow{d} \Lambda(\phi) + (1 - \gamma)^2 N_1(\phi)$, where

$$\Lambda(\phi) = \frac{\lambda \int_0^1 (x^{-\gamma} - \int_0^1 t^{-\gamma} \phi(1-t) dt) x^{-\gamma-1} \phi(1-x) \left(\int_0^x \left(\frac{y}{x}\right)^\gamma \frac{y^\rho - x^\rho}{\rho} dy \right) dx}{\int_0^1 x^{-2\gamma} \phi(1-x) dx - \left(\int_0^1 x^{-\gamma} \phi(1-x) dx \right)^2}.$$

Remark 3.1 For $\phi(x) \equiv 1$ with $0 \leq x \leq 1$, it is easy to see that $\hat{\beta}_1(\phi)$, $\hat{\gamma}(\phi)$, $\hat{\beta}_1^X(\phi)$, $\hat{\gamma}^X(\phi)$, $N_1(\phi)$ and $\Lambda(\phi)$ are the same as $\hat{\beta}_1$, $\hat{\gamma}$, $\hat{\beta}_1^X$, $\hat{\gamma}^X$, N_1 and Λ , respectively.

Example 3.1 Let's turn to Eq. 1.3 and consider the variance of Z_i conditional on $Y_{i,k}$. Simple calculation shows that for $0 \leq i < k$

$$\text{Var}(Z_i | Y_{i,k}) = \frac{(\sigma + \gamma Y_{i,k})^2}{(k-i)(1-2\gamma)(1-\gamma)^2},$$

which can be approximated by

$$\frac{1}{k} \left(1 - \frac{i}{k}\right)^{-2\gamma-1} \frac{\sigma^2}{(1-2\gamma)(1-\gamma)^2}$$

for large k since $\sup_{0 \leq i < k} |i/k - W_{\gamma,\sigma}(Y_{i,k})| \rightarrow 0$ a.s. as $k \rightarrow \infty$. Similarly to the weighted least squares estimation (i.e. $\hat{\beta} = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} Y$ in the linear model $Y = X\beta + e$ with $e \sim N(0, \Sigma)$) but without considering the correlation between $Z_i | Y_{i,k}$ and $Z_j | Y_{j,k}$, we derive the weighted least squares estimators $\hat{\beta}_1(\varphi)$ and $\hat{\gamma}(\varphi)$, where $\varphi(x) = (2\gamma + 2)(1-x)^{2\gamma+1}$ for $0 \leq x \leq 1$ and $-1 < \gamma < 0$. Consequently, we obtain $\hat{\beta}_1^X(\varphi)$ and $\hat{\gamma}^X(\varphi)$. In this case,

$$N_1(\varphi) = 2\gamma^{-2}(\gamma + 2)^2 \sigma_\gamma \int_0^1 \left(1 - \frac{2\gamma + 2}{\gamma + 2} x^\gamma\right) W(x) dx.$$

In application, since we do not know the value of γ , we can replace the weight function φ by $\hat{\varphi}$, where $\hat{\varphi}(x) = (2\hat{\gamma} + 2)(1-x)^{2\hat{\gamma}+1}$ and $\hat{\gamma}$ is some consistent estimator of γ (for example, we can let $\hat{\gamma}$ be $\hat{\gamma}^X$). Under the same conditions we can prove that $\sqrt{k}(\hat{\gamma}^X(\hat{\varphi}) - \gamma) \xrightarrow{d} \Lambda(\varphi) + (1 - \gamma)^2 N_1(\varphi)$.

3.2 Comparison

Note that N_1 is a normal random variable with mean zero, and

$$\begin{aligned} \text{Var}(N_1) &= E \left(\sigma_\gamma^{-1} \int_0^1 \left(x^{-\gamma} - \frac{1}{1-\gamma} \right) x^{-\gamma-1} W(x) dx \right)^2 \\ &= 2\sigma_\gamma^{-2} \int_0^1 \int_t^1 \left(t^{-\gamma} - \frac{1}{1-\gamma} \right) \left(s^{-\gamma} - \frac{1}{1-\gamma} \right) t^{-2\gamma} s^{-1} ds dt. \end{aligned}$$

With simple calculations (for example, by Maple version 9.0) we get

$$\text{Var}(N_1) = \frac{2(1-\gamma-12\gamma^3)}{(1-2\gamma)(1-3\gamma)^2(1-4\gamma)},$$

and if $k^{1/2}A(n/k) \rightarrow \lambda = 0$,

$$k^{1/2}(\hat{\gamma}^X - \gamma) \xrightarrow{d} N\left(0, \frac{2(1-\gamma)^4(1-\gamma-12\gamma^3)}{(1-2\gamma)(1-3\gamma)^2(1-4\gamma)}\right).$$

Again, a simple calculation shows that

$$\text{Var}(N_1(\varphi)) = \frac{16(\gamma+2)^2(1+\gamma+\gamma^2)}{3(\gamma+3)(2\gamma+3)(2-\gamma)(1-\gamma)^2(1-2\gamma)},$$

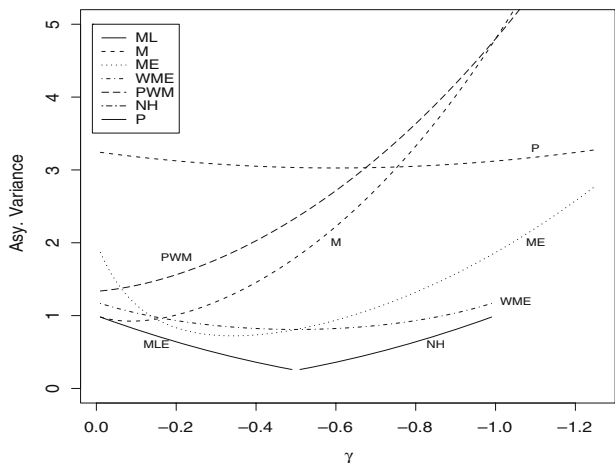
where $\varphi(x) = (2\gamma+2)(1-x)^{2\gamma+1}$, $x \in [0, 1]$, and hence if $k^{1/2}A(n/k) \rightarrow \lambda = 0$,

$$k^{1/2}(\hat{\gamma}^X(\varphi) - \gamma) \xrightarrow{d} N\left(0, \frac{16(\gamma+2)^2(1-\gamma)^2(1+\gamma+\gamma^2)}{3(\gamma+3)(2\gamma+3)(2-\gamma)(1-2\gamma)}\right).$$

One interesting fact is that the asymptotic variance of $\hat{\gamma}^X(\varphi)$ is symmetric about $\gamma = -1/2$ and attains its smallest value 0.81 at $\gamma = -1/2$ (see Fig. 1 below).

Until now there exist five well-known estimators for the negative extreme value index: maximum likelihood (ML) estimator (Drees et al. 2004; Smith 1985), moment (M) estimator (Dekkers et al. 1989), Pickands (P) estimator (Pickands 1975), probability weighted moment (PWM) estimator (Hosking et al. 1985) and negative Hill (NH) estimator (Falk 1995). Our new mean excess (ME) estimator and weighted mean excess (WME) estimator related to φ are scale and shift invariant as the mentioned five estimators. We plot the asymptotic variance of all the estimators for $\gamma \in (-1.25, 0)$ if possible. From Fig. 1, we see that the maximum likelihood estimator and the negative Hill estimator have the smallest asymptotic variance and that our new estimators are of smaller asymptotic variance for most of γ . Thus, our mean excess estimator and weighted mean excess estimator are comparable with the other estimators. Note that the asymptotic normality results for the maximum likelihood estimator and the negative Hill estimator only hold for $-1/2 < \gamma < 0$ and $-1 < \gamma < -1/2$, respectively, while the weighted mean excess estimator (related

Fig. 1 Asymptotic variance of the estimators: maximum likelihood (ML), moment (M), Pickands (P), negative Hill (NH), probability weighted moment (PWM) estimators, mean excess (ME) estimator and weighted mean excess (WME) estimator related to φ



to φ) holds for $-1 < \gamma < 0$ and the others hold for $\gamma < 0$. Since the maximum likelihood estimator is obtained by numeric calculating the roots for two equations, its computation capacity is very large and sometimes the roots can not be obtained (see Hüsler and Li 2006). Meanwhile, our (weighted) mean excess estimators are explicit functions of the order statistics and can be easily calculated. In view of these points, we believe that our (weighted) mean excess estimators are good replacements for the maximum likelihood estimator and negative Hill estimator for $\gamma < 0$.

4 Proofs

Proof of Theorem 2.1 Note that $F(t) = P(U_1 \leq t) = t^\alpha$. Let f and Q be the density function and the inverse function of F , respectively. Then $f(t) = \alpha t^{\alpha-1}$ and $Q(t) = t^{1/\alpha}$. Let

$$\rho_k(t) = k^{1/2} f(Q(t))(U_{[kt],k} - Q(t)),$$

where $[kt]$ denotes the smallest integer not less than kt . Then by the weighted approximations to the general quantile process (e.g. see Theorem 2.1 in Csörgő and Horváth 1993, page 381), there exist a sequence of Brownian bridges $\{B_k(t), t \in [0, 1]\}$ such that for $0 < v \leq 1/2$

$$k^{1/2-v} \sup_{\frac{1}{k+1} \leq t \leq \frac{k}{k+1}} t^{-v} |\rho_k(t) - B_k(t)| = O_P(1)$$

as $k \rightarrow \infty$. Thus, for $i = 1, 2, \dots, k - 1$, we obtain that

$$\begin{aligned} U_{i,k} &= Q(i/k) + k^{-1/2} (f(Q(i/k)))^{-1} [B_k(i/k) + k^{-1/2+v} (i/k)^v O_P(1)] \\ &= (i/k)^{1/\alpha} \left(1 + \alpha^{-1} k^{-1/2} (i/k)^{-1} B_k(i/k) + \alpha^{-1} k^{-1+v} (i/k)^{-1+v} O_P(1) \right) \\ &=: (i/k)^{1/\alpha} (1 + \Delta(k, i)) \end{aligned} \tag{4.1}$$

for sufficiently large k , where the O_P -item is uniform in i .

Without loss of generality we assume $B_k(t) = W_k(t) - tW_k(1)$, where W_k is a standard Brownian motion. By the laws of the iterated logarithm for Brownian motion (see e.g. Borodin and Salminen 2002, page 53), it follows that for each ε such that $0 < \varepsilon < 2/\alpha$

$$\sup_{\{i: 1 \leq i \leq k\}} (i/k)^{-1/2+\varepsilon/2} |B_k(i/k)| = O_P(1).$$

Hence

$$\sup_{\{i: 1 \leq i \leq k\}} |U_{i,k} - (i/k)^{1/\alpha}| = o_P(1) \tag{4.2}$$

and

$$\sup_{\{i: k^{1-1/(1+2\varepsilon)} \leq i \leq k-1\}} \Delta(k, i) = o_P(1). \tag{4.3}$$

By Eq. 4.1 we have

$$\begin{aligned} \sum_{j=1}^{i-1} U_{j,k} &= \sum_{j=1}^{i-1} (j/k)^{1/\alpha} + \alpha^{-1} k^{-1/2} \sum_{j=1}^{i-1} (j/k)^{-1+1/\alpha} B_k(j/k) \\ &\quad + \alpha^{-1} k^{-1+v} O_P(1) \sum_{j=1}^{i-1} (j/k)^{-1+1/\alpha+v} \end{aligned}$$

and

$$\begin{aligned} &k^{-1/2} \sum_{j=1}^{i-1} \left(\frac{U_{j,k}}{U_{i,k}} - \frac{\alpha}{1+\alpha} \right) \\ &= \frac{k^{-1/2} \sum_{i=1}^{i-1} [(j/k)^{1/\alpha} - \frac{\alpha}{1+\alpha} (i/k)^{1/\alpha}]}{(i/k)^{1/\alpha} (1 + \Delta(k, i))} - \frac{(i-1)k^{-1/2} \frac{\alpha}{1+\alpha} (i/k)^{1/\alpha} \Delta(k, i)}{(i/k)^{1/\alpha} (1 + \Delta(k, i))} \\ &\quad + \frac{\alpha^{-1} k^{-1} \sum_{j=1}^{i-1} (j/k)^{-1+1/\alpha} B_k(j/k)}{(i/k)^{1/\alpha} (1 + \Delta(k, i))} + \frac{\alpha^{-1} k^{-1/2+v} O_P(1) k^{-1} \sum_{j=1}^{i-1} (j/k)^{-1+1/\alpha+v}}{(i/k)^{1/\alpha} (1 + \Delta(k, i))} \\ &=: \tau_1(k, i) - \tau_2(k, i) + \tau_3(k, i) + \tau_4(k, i). \end{aligned}$$

We first show that

$$\begin{aligned} &\sup_{k^{-1/(1+2\epsilon)} \leq t < 1} \left| k^{-1/2} \sum_{j=1}^{\lfloor kt \rfloor - 1} \left(\frac{U_{j,k}}{U_{\lfloor kt \rfloor, k}} - \frac{\alpha}{1+\alpha} \right) \right. \\ &\quad \left. - \frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B_k(u) du + \frac{1}{1+\alpha} B_k(t) \right| = o_P(1). \end{aligned} \tag{4.4}$$

In order to show Eq. 4.4, it suffices to verify that as $k \rightarrow \infty$

$$\sup_{\{i: k^{1-1/(1+2\epsilon)} \leq i \leq k-1\}} \sup_{\{t: i/k \leq t < (i+1)/k\}} |\tau_1(k, i)| = o_P(1), \tag{4.5}$$

$$\sup_{\{i: k^{1-1/(1+2\epsilon)} \leq i \leq k-1\}} \sup_{\{t: i/k \leq t < (i+1)/k\}} \left| \tau_2(k, i) - \frac{1}{1+\alpha} B_k(t) \right| = o_P(1), \tag{4.6}$$

$$\sup_{\{i: k^{1-1/(1+2\epsilon)} \leq i \leq k-1\}} \sup_{\{t: i/k \leq t < (i+1)/k\}} \left| \tau_3(k, i) - \frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B_k(u) du \right| = o_P(1), \tag{4.7}$$

and

$$\sup_{\{i: k^{1-1/(1+2\epsilon)} \leq i \leq k-1\}} \sup_{\{t: i/k \leq t < (i+1)/k\}} |\tau_4(k, i)| = o_P(1). \tag{4.8}$$

Relations (4.5) and (4.8) hold immediately by Eq. 4.3. By the modulus of continuity for Brownian motion (see e.g. Borodin and Salminen 2002, page 53) it follows that

$$\sup_{i/k \leq t < (i+1)/k} |B_k(i/k) - B_k(t)| = O_P((k \log k)^{-1/2}). \tag{4.9}$$

Using Eqs. 4.3 and 4.9, we can show that Eqs. 4.6 and 4.7 follow. Thus Eq. 4.4 holds.

On the other hand, for $0 < \varepsilon < 1/2$

$$\sup_{0 < t < k^{-1/(1+2\varepsilon)}} \left| k^{-1/2} \sum_{j=1}^{[kt]-1} \left(\frac{U_{j,k}}{U_{[kt],k}} - \frac{\alpha}{1+\alpha} \right) \right| \leq k^{-1/2} 2k^{1-1/(1+2\varepsilon)} = o(1) \quad a.s.$$

and

$$\begin{aligned} & \sup_{0 < t < k^{-1/(1+2\varepsilon)}} \left| \frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B_k(u) du - \frac{1}{1+\alpha} B_k(t) \right| \\ & \leq \sup_{0 < t < k^{-1/(1+2\varepsilon)}} |B_k(t)| \sup_{0 < t < k^{-1/(1+2\varepsilon)}} \frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} du + \sup_{0 < t < k^{-1/(1+2\varepsilon)}} \frac{1}{1+\alpha} |B_k(t)| \\ & \leq \sup_{0 < t < k^{-1/(1+2\varepsilon)}} 2|B_k(t)| \\ & \leq \sup_{0 < t < k^{-1/(1+2\varepsilon)}} 2(|W_k(t)| + t|W_k(1)|) = o_P(1). \end{aligned}$$

So,

$$\sup_{0 < t < 1} \left| k^{-1/2} \sum_{j=1}^{[kt]-1} \left(\frac{U_{j,k}}{U_{[kt],k}} - \frac{\alpha}{1+\alpha} \right) - \left(\frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B_k(u) du - \frac{1}{1+\alpha} B_k(t) \right) \right| = o_P(1)$$

and the statement of Theorem 2.1 follows by the fact

$$\frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} B_k(u) du - \frac{1}{1+\alpha} B_k(t) = \frac{1}{\alpha t^{1/\alpha}} \int_0^t u^{-1+1/\alpha} W_k(u) du - \frac{1}{1+\alpha} W_k(t).$$

□

By Theorem 2.1 and the Skorohod construction, there exist a sequence of Gaussian processes $\{W_k(t) : t \in [0, 1]\}$, satisfying that for each $k \geq 1$ and $0 \leq t \leq s \leq 1$, $EW_k(t) = 0$ and $EW_k(t)W_k(s) = t(t/s)^{-\gamma}$, such that

$$\sup_{t \in [0, 1]} |T_k(t) - W_k(t)| \xrightarrow{P} 0, \quad \text{as } k \rightarrow \infty. \tag{4.10}$$

In the proofs of Theorem 2.2 and 2.3, we will use these Gaussian processes W_k 's of Eq. 4.10.

Proof of Theorem 2.2 Note that for $i \geq 2$,

$$\begin{aligned} \frac{V_i}{U_{i,k}} &= \frac{1}{i-1} \sum_{j=1}^{i-1} \frac{U_{j,k}}{U_{i,k}} - 1 = \sigma_\gamma^{-1} k^{-1/2} \sum_{j=1}^{i-1} \left(\frac{U_{j,k}}{U_{i,k}} - \frac{1}{1-\gamma} \right) \cdot k^{1/2} (i-1)^{-1} \sigma_\gamma + \frac{\gamma}{1-\gamma} \\ &= (W_k(i/k) + o_P(1)) k^{1/2} (i-1)^{-1} \sigma_\gamma + \frac{\gamma}{1-\gamma}, \quad (\text{by Eq. (4.10)}) \end{aligned}$$

where $o_P(1)$ is uniform in i . Denote

$$\tau(i, k) = U_{i,k} (W_k(i/k) + o_P(1)) k^{1/2} (i-1)^{-1} \sigma_\gamma$$

and $\bar{\tau} = \frac{1}{k} \sum_{i=2}^{k+1} \tau(i, k)$. Then $V_i = \tau(i, k) + \gamma(1-\gamma)^{-1}U_{i,k}$ and $\bar{V} = \bar{\tau} + \gamma(1-\gamma)^{-1}\bar{U}$. Since

$$\begin{aligned} \frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})(V_i - \bar{V}) &= \frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U}) \left(\frac{\gamma}{1-\gamma} (U_{i,k} - \bar{U}) + \tau(i, k) - \bar{\tau} \right) \\ &= \frac{\gamma}{1-\gamma} \frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})^2 + \frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})\tau(i, k), \end{aligned}$$

it follows that

$$\hat{\beta}_1 = \frac{\gamma}{1-\gamma} + \frac{\frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})\tau(i, k)}{\frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})^2}. \tag{4.11}$$

Recall that, with $\beta_1 = \gamma/(1 + \gamma)$ and Eq. 4.11,

$$k^{1/2}(\hat{\beta}_1 - \beta_1) = \frac{\frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})k^{1/2}\tau(i, k)}{\frac{1}{k} \sum_{i=2}^{k+1} (U_{i,k} - \bar{U})^2}.$$

By Eq. 4.2, it follows that

$$k^{1/2}\tau(i, k) = \sigma_\gamma(i/k)^{-\gamma-1}W_k(i/k) + o_P(1)$$

and

$$k^{1/2}\bar{\tau} = \sigma_\gamma \frac{1}{k} \sum_{i=2}^k (i/k)^{-\gamma-1}W_k(i/k) + o_P(1)$$

for large k , where $o_P(1)$ is uniform in i . Again, by Eq. 4.2,

$$\begin{aligned} k^{1/2}(\hat{\beta}_1 - \beta_1) &= \frac{\sigma_\gamma \frac{1}{k} \sum_{i=2}^{k+1} ((i/k)^{-\gamma} - \frac{1}{1-\gamma})(i/k)^{-\gamma-1}W_k(i/k) + o_P(1)}{\sigma_\gamma^2 + o_P(1)} \\ &= \sigma_\gamma^{-1} \int_0^1 \left(x^{-\gamma} - \frac{1}{1-\gamma} \right) x^{-\gamma-1}W_k(x)dx + o_P(1). \end{aligned}$$

On the other hand, by Eq. 2.3,

$$\begin{aligned} \sqrt{k}(\hat{\beta}_0 - \beta_0) &= \frac{\sigma}{\gamma} \sqrt{k}(\bar{V} - \hat{\beta}_1(\bar{U} - 1) - \beta_1) \\ &= \frac{\sigma}{\gamma} \left(\sqrt{k}\bar{\tau}(k) + \sqrt{k}(\hat{\beta}_1 - \beta_1)(1 - \bar{U}) \right) \\ &= \frac{\sigma}{\gamma} \left(\sigma_\gamma \int_0^1 x^{-\gamma-1}W_k(x)dx - \frac{\gamma}{1-\gamma} \sigma_\gamma^{-1} \int_0^1 \left(x^{-\gamma} - \frac{1}{1-\gamma} \right) x^{-\gamma-1}W_k(x)dx \right) \\ &\quad + o_P(1). \end{aligned}$$

Thus Eq. 2.6 holds.

Now consider $\hat{\gamma}$ and $\hat{\sigma}$. By Eq. 1.5 we have

$$k^{1/2}(\hat{\gamma} - \gamma) = k^{1/2} \left(\frac{\hat{\beta}_1}{1 + \hat{\beta}_1} - \frac{\beta_1}{1 + \beta_1} \right) = \frac{k^{1/2} (\hat{\beta}_1 - \beta_1)}{(1 + \hat{\beta}_1)(1 + \beta_1)}$$

and

$$k^{1/2}(\hat{\sigma} - \sigma) = k^{1/2} \left(\frac{\hat{\beta}_0}{1 + \hat{\beta}_1} - \frac{\beta_0}{1 + \beta_1} \right) = \frac{k^{1/2} (\hat{\beta}_0 - \beta_0)}{1 + \hat{\beta}_1} - \frac{\beta_0 k^{1/2} (\hat{\beta}_1 - \beta_1)}{(1 + \hat{\beta}_1)(1 + \beta_1)}.$$

Thus Eq. 2.7 immediately follows by Eq. 2.6 and so Theorem 2.2 holds. □

Proof of Theorem 2.3 Let E_1, E_2, \dots, E_n be i.i.d. random variables distributed with $F_E(x) = 1 - 1/x, x \geq 1$. Then

$$\{X_{n-i,n}\}_{i=0}^k \stackrel{d}{=} \{U(E_{n-i,n})\}_{i=0}^k,$$

where $U(t) = F^{\leftarrow}(1 - 1/t)$ for $t \geq 1$. It means that

$$\begin{aligned} \sum_{i=0}^{k-1} (\tilde{Y}_{i,k} - \bar{Y}) &= \sum_{i=0}^{k-1} \left(X_{n-k+i,n} - \frac{1}{k} \sum_{j=0}^{k-1} X_{n-k+j,n} \right) \\ &\stackrel{d}{=} \sum_{i=0}^{k-1} \left(U(E_{n-k+i,n}) - \frac{1}{k} \sum_{j=0}^{k-1} U(E_{n-k+j,n}) \right) \\ &= \sum_{i=0}^{k-1} \left(U_{\infty}(E_{n-k+i,n}) - \frac{1}{k} \sum_{j=0}^{k-1} U_{\infty}(E_{n-k+j,n}) \right) \times [U(E_{n-k,n}) - U(\infty)] \end{aligned}$$

where

$$U_{\infty}(E_{n-k+i,n}) = \frac{U(\infty) - U(E_{n-k+i,n})}{U(\infty) - U(E_{n-k,n})}.$$

Defining

$$\tilde{U}_{\infty}(n, k) = \frac{1}{k} \sum_{i=0}^{k-1} U_{\infty}(E_{n-k+i,n}),$$

$$V_{\infty}(E_{n-k+i,n}) = \frac{1}{k-i} \sum_{j=i+1}^k U_{\infty}(E_{n-k+j,n}) - U_{\infty}(E_{n-k+i,n})$$

and

$$\tilde{V}_{\infty}(n, k) = \frac{1}{k} \sum_{i=0}^{k-1} V_{\infty}(E_{n-k+i,n}),$$

we can also rewrite the denominator and the second factor in the numerator of $\hat{\beta}_1^X$. Thus $\hat{\beta}_1^X \stackrel{d}{=} \hat{\beta}_1^E$, where

$$\hat{\beta}_1^E = \frac{\sum_{i=0}^{k-1} (U_\infty(E_{n-k+i,n}) - \bar{U}_\infty(n, k))(V_\infty(E_{n-k+i,n}) - \bar{V}_\infty(n, k))}{\sum_{i=0}^{k-1} (U_\infty(E_{n-k+i,n}) - \bar{U}_\infty(n, k))^2}$$

Note that $E^\gamma \sim U_\alpha[0, 1]$ with $\alpha = -1/\gamma > 0$, since $P(E^\gamma \leq x) = P(E \geq x^{1/\gamma}) = x^{-1/\gamma}$ for $x \in [0, 1]$. Hence

$$\left\{ \left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}} \right)^\gamma \right\}_{0 \leq i < j \leq k} \stackrel{d}{=} \left\{ \frac{U_{k-j+1,n}}{U_{k-i+1,n}} \right\}_{0 \leq i < j \leq k} \stackrel{d}{=} \left\{ \frac{U_{k-j+1,k}}{U_{k-i+1,k}} \right\}_{0 \leq i < j \leq k}, \tag{4.12}$$

where U_1, U_2, \dots are i.i.d. random variables distributed with $U_\alpha[0, 1]$.

By Theorem 2.1, there exists a sequence of Gaussian processes W_k 's with mean zero such that $EW_k(t)W_k(s) = t(t/s)^{-\gamma}$ for $0 \leq t \leq s \leq 1$ and that by Eq. 4.10

$$\frac{1}{k-i} \sum_{j=i+1}^k \left(\left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}} \right)^\gamma - \frac{1}{1-\gamma} \right) = \frac{k^{1/2}\sigma_\gamma}{k-i+1} \left(W_k \left(\frac{k-i+1}{k} \right) + o_P(1) \right). \tag{4.13}$$

The o_P -term above is uniform for all $0 \leq i < j \leq k$.

By Eq. 2.8 and Lemma 2.1 in Drees (1998), it follows that for each $\varepsilon > 0$, there exist $A_0 \sim A, t_0 > 0$ such that for all $t \geq t_0$ and $x \geq 1$,

$$\left| \frac{\frac{U(\infty)-U(tx)}{U(\infty)-U(t)} - x^\gamma}{A_0(t)} - x^\gamma \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon x^{\gamma+\rho+\varepsilon}.$$

Without loss of generality, we assume $A_0 = A$. First consider the case $\rho < 0$. Let $0 < \varepsilon < -\rho$. Then for large k , with $x = E_{n-k+i,n}/E_{n-k,n}$ and $t = E_{n-k,n}$

$$U_\infty(E_{n-k+i,n}) = \left(\frac{E_{n-k+i,n}}{E_{n-k,n}} \right)^\gamma \left(1 + A(E_{n-k,n}) \times \left[\frac{\left(\frac{E_{n-k+i,n}}{E_{n-k,n}} \right)^\rho - 1}{\rho} + o_P(1) \left(\frac{E_{n-k+i,n}}{E_{n-k,n}} \right)^{\rho+\varepsilon} \right] \right) \tag{4.14}$$

and the $o_P(1)$ -term is uniform for $0 \leq i \leq k$. Hence

$$\sup_{0 \leq i \leq k} \left| U_\infty(E_{n-k+i,n}) - \left(\frac{E_{n-k+i,n}}{E_{n-k,n}} \right)^\gamma \right| = o_P(1) \tag{4.15}$$

holds since $A(t) \rightarrow 0$ and

$$\frac{U_\infty(E_{n-k+j,n})}{U_\infty(E_{n-k+i,n})} = \left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}} \right)^\gamma \left(1 + A(E_{n-k,n}) \left(\frac{E_{n-k+j,n}^\rho - E_{n-k+i,n}^\rho}{\rho E_{n-k,n}^\rho} + o_P(1) \right) \right)$$

follows for large k and the $o_P(1)$ -term is uniform for $0 \leq i < j \leq k$. Hence using Eq. 4.13

$$\begin{aligned} & \frac{1}{k-i} \sum_{j=i+1}^k \left(\frac{U_\infty(E_{n-k+j,n})}{U_\infty(E_{n-k+i,n})} - \frac{1}{1-\gamma} \right) \\ &= \frac{1}{k-i} \sum_{j=i+1}^k \left(\left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}} \right)^\gamma - \frac{1}{1-\gamma} \right) \\ & \quad + \frac{A(E_{n-k,n})}{k-i} \sum_{j=i+1}^k \left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}} \right)^\gamma \left(\frac{E_{n-k+j,n}^\rho - E_{n-k+i,n}^\rho}{\rho E_{n-k,n}^\rho} + o_P(1) \right) \\ &= \frac{k^{1/2} \sigma_\gamma}{k-i+1} \left(W_k \left(\frac{k-i+1}{k} \right) + o_P(1) \right) \\ & \quad + \frac{A(E_{n-k,n})}{k-i} \sum_{j=i+1}^k \left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}} \right)^\gamma \left(\frac{E_{n-k+j,n}^\rho - E_{n-k+i,n}^\rho}{\rho E_{n-k,n}^\rho} + o_P(1) \right) \\ &=: \frac{\tau(i, k)}{U_\infty(E_{n-k+i,n})}. \end{aligned}$$

So,

$$V_\infty(E_{n-k+i,n}) = \frac{\gamma}{1-\gamma} U_\infty(E_{n-k+i,n}) + \tau(i, k)$$

and, with notation $\bar{\tau} = k^{-1} \sum_{i=0}^{k-1} \tau(i, k)$,

$$\bar{V}_\infty(k, n) = \frac{\gamma}{1-\gamma} \bar{U}_\infty(n, k) + \bar{\tau}.$$

Similar to the proof of Theorem 2.2, we have

$$k^{1/2}(\hat{\beta}_1^E - \beta_1) = \frac{\frac{1}{k} \sum_{i=0}^{k-1} k^{1/2} \tau(i, k) (U_\infty(E_{n-k+i,n}) - \bar{U}_\infty(n, k))}{\frac{1}{k} \sum_{i=0}^{k-1} (U_\infty(E_{n-k+i,n}) - \bar{U}_\infty(n, k))^2}.$$

By Eq. 4.12,

$$\{(E_{n-k+i,n}/E_{n-k,n})^\gamma\}_{0 \leq i \leq k} \stackrel{d}{=} \{U_{k-i+1,k}\}_{0 \leq i \leq k}$$

and by Eqs. 4.2 and 4.15, we get

$$\sup_{0 \leq i \leq k} |U_\infty(E_{n-k+i,n}) - ((k-i)/k)^{-\gamma}| = o_P(1).$$

Again, similar to the proof of Theorem 2.2, it follows that

$$\begin{aligned} k^{1/2} \tau(i, k) &= \sigma_\gamma \left(\frac{k-i}{k} \right)^{-\gamma-1} W_k \left(\frac{k-i}{k} \right) \\ & \quad + \frac{\lambda}{\rho} \left(\frac{k-i}{k} \right)^{-\gamma-1} \frac{1}{k} \sum_{j=i+1}^k \left(\frac{k-j}{k-i} \right)^\gamma \left(\left(\frac{k-j}{k} \right)^\rho - \left(\frac{k-i}{k} \right)^\rho \right) + o_P(1), \end{aligned}$$

since $(k/n)E_{n-k,n} \xrightarrow{P} 1$ and $k^{1/2}A(n/k) \rightarrow \lambda$. Hence

$$k^{1/2}(\hat{\beta}_1^X - \beta_1) \stackrel{d}{=} k^{1/2}(\hat{\beta}_1^E - \beta_1) \xrightarrow{d} \sigma_\gamma^{-1} \int_0^1 \left(x^{-\gamma} - \frac{1}{1-\gamma}\right) x^{-\gamma-1} W(x) dx + \lambda \sigma_\gamma^{-2} \int_0^1 \left(x^{-\gamma} - \frac{1}{1-\gamma}\right) x^{-\gamma-1} \int_0^x \left(\frac{y}{x}\right)^\gamma \frac{y^\rho - x^\rho}{\rho} dy dx.$$

For $\rho = 0$, let ε be such that $0 < \varepsilon < \min\{-\gamma, 1/2\}$, then Eq. 4.14 and 4.15 still hold with the replacement $\log(E_{n-k+i,n}/E_{n-k,n})$ for $\left(\left(\frac{E_{n-k+i,n}}{E_{n-k,n}}\right)^\rho - 1\right)/\rho$. Since $k^{-1/2}U_{1,k}^{\varepsilon/\gamma} = k^{-1/2+\varepsilon} \left(kU_{1,k}^{-1/\gamma}\right)^{-\varepsilon} = o_P(1)$, it follows that for all $0 \leq i \leq k$

$$\left|A(E_{n-k,n}) \left(\frac{E_{n-k+i,n}}{E_{n-k,n}}\right)^\varepsilon\right| \stackrel{d}{=} \left|k^{1/2}A(E_{n-k,n})(k^{-1/2}U_{k-i+1,k}^{\varepsilon/\gamma})\right| = o_P(1).$$

Thus, by Taylor expansion, Eq. 4.14 implies

$$\frac{U_\infty(E_{n-k+j,n})}{U_\infty(E_{n-k+i,n})} = \left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}}\right)^\gamma \left(1 + A(E_{n-k,n}) \log\left(\frac{E_{n-k+j,n}}{E_{n-k+i,n}}\right) + o_P(k^{-1/2})\right).$$

The rest of the proof follows the steps of the Proof of Theorem 2.2.

The asymptotic normality of $\hat{\gamma}^X$ follows by that of $\hat{\beta}_1^X$. Thus Theorem 2.3 holds. □

Acknowledgements The authors thank the associate editor and referees for valuable comments. Part of the second author’s work was done when he worked at University of Bern, supported by a grant of the Swiss National Science foundation.

References

J. Beirlant, J. L. Teugels, and P. Vynckier, *Practical Analysis of Extreme Values*, Leuven University Press, 1996.

A. N. Borodin and P. Salminen, *Handbook of Brownian Motion – Facts and Formulae*, Second Edition. Birkhäuser, 2002.

M. Csörgő and L. Horváth, *Weighted Approximations in Probability and Statistics*, Wiley, 1993.

A. L. M. Dekkers, L. de Haan, and J. H. J. Einmahl, “A moment estimator for the index of an extreme-value distribution,” *Annals of Statistics* vol. 17 pp. 1833–1855, 1989.

H. Drees, “On smooth statistical tail functionals,” *Scandinavian Journal of Statistics* vol. 25, pp. 187–210, 1998.

H. Drees, A. Ferreira, and L. de Haan, “On the maximum likelihood estimation of the extreme value index,” *Annals of Applied Probability* vol. 14, pp. 1179–1201, 2004.

M. Falk, “Some best parameter estimates for distributions with finite endpoint,” *Statistics* vol. 27, pp. 115–125, 1995.

L. de Haan and U. Stadtmüller, “Generalized regular variation of second order,” *Journal of the Australian Mathematical Society. Series A* vol. 61, pp. 381–395, 1996.

J. R. M. Hosking, J. R. Wallis, and E. F. Wood, “Estimation of the generalized extreme-value distribution by the method of probability-weighted moments,” *Technometrics* vol. 27, pp. 251–261, 1985.

J. Hüslér and D. Li, “On testing extreme value conditions,” *Extremes* vol. 9, pp. 69–86, 2006.

J. Pickands, “Statistical inference using extreme order statistics,” *Annals of Statistics* vol. 3, pp. 119–131, 1975.

R. D. Reiss, and M. Thomas, *Statistical Analysis of Extreme Values*, 3rd Edition. Birkhäuser, Basel, 2007.

S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, 1987

R. L. Smith, “Maximum likelihood estimation in a class of nonregular cases,” *Biometrika* vol. 72, pp. 67–90, 1985