Regression estimators for the tail index

Amenah AL-Najafi, László Viharos

Bolyai Institute, University of Szeged, Aradi vėrtanák tere 1, Szeged, H-6720, Hungary

Abstract

We propose a class of weighted least squares estimators for the tail index of a distribution function with a regularly varying upper tail. Our approach is based on the method developed by Holan and McElroy (2010) for the Parzen tail index. Asymptotic normality of the estimators is proved. Through a simulation study, these and earlier estimators are compared in the Pareto and Hall models using the mean squared error as criterion. The results show that the weighted least squares estimator is better than the other estimators investigated.

Keywords: tail index, weighted least squares estimators, Pareto model, quantile process.

1. Introduction and main result

Let \( X_1, X_2, \ldots \) be independent random variables with a common right-continuous distribution function \( F \), and for each \( n \in \mathbb{N} \), let \( X_{1,n} \leq \cdots \leq X_{n,n} \) denote the order statistics pertaining to the sample \( X_1, \ldots, X_n \). Let \( \mathcal{R}_\alpha \) be the class of all distribution functions \( F \) such that \( 1 - F \) is regularly varying at infinity with index \(-1/\alpha\), that is,

\[
1 - F(x) = x^{-1/\alpha} \ell(x), \quad 1 < x < \infty,
\]

where \( \ell \) is some positive function on the half line \([1, \infty)\), slowly varying at infinity and \( \alpha > 0 \) is a fixed unknown parameter to be estimated. Introducing the quantile function \( Q \) of \( F \) defined as

\[
Q(s) := \inf \{ x : F(x) \geq s \}, \quad 0 < s \leq 1, \quad Q(0) := Q(0+),
\]

corresponding author.

Email addresses: amenah@math.u-szeged.hu (Amenah AL-Najafi), viharos@math.u-szeged.hu (László Viharos)
it is well known that $F \in \mathcal{R}_\alpha$ if and only for some function $L$ slowly varying at zero,

$$Q(1 - s) = s^{-\alpha}L(s), \quad 0 < s < 1. \quad (1)$$

Several estimators exist for the tail index $\alpha$ among which Hill’s estimator is the most classical. Hill (1975) proposed the following estimator for the tail index $\alpha$:

$$\hat{\alpha}_n(H) = \frac{1}{k_n} \sum_{j=1}^{k_n} \log X_{n-j+1,n} - \log X_{n-k_n,n},$$

where the $k_n$ are positive integers, which in theoretical asymptotic considerations will satisfy the conditions

$$1 \leq k_n < n, \quad k_n \to \infty \quad \text{and} \quad k_n/n \to 0 \quad \text{as} \quad n \to \infty.$$ 

The asymptotic normality of $\hat{\alpha}_n(H)$ was first considered by Hall (1982) in the following submodel of $\mathcal{R}_\alpha$:

$$1 - F(x) = x^{-1/\alpha}C_1[1 + C_2x^{-\beta/\alpha}\{1 + o(1)\}], \quad \text{as} \quad x \to \infty,$$

for some constants $C_1 > 0$ and $C_2 \neq 0$. This is equivalent to

$$Q(1 - s) = s^{-\alpha}D_1[1 + D_2s^\beta\{1 + o(1)\}], \quad s \to 0, \quad (2)$$

where $D_1 = C_1^\alpha$ and $D_2 = C_2/C_1^\beta$.

Another estimators were proposed by Pickands (1975), Dekkers et al. (1989), to name a few.

Assuming that $F$ is absolutely continuous with density function $f$, Parzen (2004) studied the following alternative model for the right tail of the distribution:

$$fQ(s) := f(Q(s)) = (1 - s)^\nu L_1(1 - s), \quad s \in (1/2, 1],$$

where $\nu > 0$ is a finite constant and $L_1$ is slowly varying at zero. The parameter $\nu$ is called the Parzen tail index of the density-quantile function $fQ(\cdot)$.

Based on an orthogonal series expansion for $L_1$, Holan and McElroy (2010) introduced a regression estimator for the Parzen tail index using ordinary least squares. AL-Najafi and Viharos (2020) obtained a more general class of estimators for $\nu$ using weighted least squares. We adopt this method to
estimate the classical tail index $\alpha$. Following the idea of Holan and McElroy (2010), we assume that the slowly varying function $L$ in (1) admits the truncated orthogonal series expansion

$$
L(s) = \exp \left\{ \theta_0 + 2 \sum_{k=1}^{p} \theta_k \cos(2\pi ks) \right\},
$$

where $p > 0$ is a fixed unknown integer, and $\theta_0, \ldots, \theta_p$ are unknown parameters. It follows that

$$
\log Q(1-s) = -\alpha \log s + \theta_0 + 2 \sum_{k=1}^{p} \theta_k \cos(2\pi ks).
$$

Let $Q_n$ be the empirical quantile function defined as

$$
Q_n(s) = X_{k,n} \quad \text{if} \quad \frac{k-1}{n} < s \leq \frac{k}{n}, \quad k = 1, 2, \ldots, n.
$$

Based on the representation (3), we obtain the regression equations

$$
\log Q_n(1-s_j) = -\alpha \log s_j + \theta_0 + 2 \sum_{k=1}^{\tilde{p}} \theta_k \cos(2\pi ks_j) + \varepsilon(s_j),
$$

where $\varepsilon(s) = \log(Q_n(1-s)/Q(1-s))$ is the residual process, $s_j = j/n$, $j = [na], \ldots, [nb]$, $a < b$ are fixed constants taken from the interval $(0,1)$, $\tilde{p} > p$ is chosen by the statistician and $\theta_k = 0$ for $k > p$. We propose a class of estimators for $\alpha$ using weighted least squares. We choose some nonnegative weights of the form $w_{j,n} = R(j/n)$ with some weight function $R$. Set $y_j := \log Q_n(1-s_j)$,

$$
y := (y_{[na]}, \ldots, y_{[nb]})',
$$

$$
W := \text{diag}(w_{[na],n}, \ldots, w_{[nb],n}),
$$

and let $X := [G^*, G_0, 2G_1, \ldots, 2G_{\tilde{p}}]$, where

$$
G^* = (-\log(s_{[na]}), \ldots, -\log(s_{[nb]}))',
$$

$$
G_k = (\cos(2\pi ks_{[na]}), \ldots, \cos(2\pi ks_{[nb]}))', \quad k = 0, \ldots, \tilde{p}.
$$

Set $\beta_{\tilde{p}} := (\alpha, \theta_0, \theta_1, \ldots, \theta_{\tilde{p}})'$. By minimizing the weighted sum of squares

$$
\sum_{[na]}^{[nb]} w_{j,n}(y_j + \alpha \log s_j - \theta_0 - 2 \sum_{k=1}^{\tilde{p}} \theta_k \cos(2\pi ks_j))^2,
$$

3
we obtain the following estimator of $\beta_{\tilde{p}}$:

$$\hat{\beta}_{\tilde{p}} = (X'W X)^{-1} X'W y.$$ 

Then the weighted least squares estimator of $\alpha$ can be written in the form

$$\hat{\alpha}^{(W)} := e_1' \hat{\beta}_{\tilde{p}} = e_1' (X'W X)^{-1} X'W y,$$

where $e_1$ is the $\tilde{p} + 2$ dimensional vector defined as $e_1 = (1, 0, 0, \ldots, 0)'$.

We assume the following conditions on the underlying distribution:

1. **(Q1)** The distribution function $F$ is continuous and twice differentiable on $(a^*, b^*)$, where $a^* = \sup \{x : F(x) = 0\}$, $b^* = \inf \{x : F(x) = 1\}$, $-\infty \leq a^* < b^* \leq \infty$ and $f(x) := F'(x) \neq 0$ on $(a^*, b^*)$.

2. **(Q2)** $\sup_{a^* < x < b^*} F(x) |1 - F(x)| f'(x)/f^2(x) | < \infty$.

3. **(Q3)** $\sup_{1 - b \leq s \leq 1 - a} 1/|Q(s)| < \infty$, $\sup_{1 - b \leq s \leq 1 - a} 1/fQ(s) < \infty$ and $\sup_{1 - b \leq s \leq 1 - a} 1/|fQ(s)Q(s)| < \infty$.

We will show that the limit matrix $M(a, b, R) := \lim_{n \to \infty} n^{-1} X'W X$ exists (see the proof of Theorem 1 in Chapter 3). Let $(v^*, v_0, \ldots, v_{\tilde{p}})$ be the first row of $M(a, b, R)^{-1}$, and set $G_R(u) := R(u)(-v^* \log u + v_0 + 2 \sum_{k=1}^{\tilde{p}} v_k \cos(2\pi ku))$ for $u \in (0, 1)$.

Moreover, we suppose the following conditions:

1. **(R)** The weight function $R$ is nonnegative and Riemann integrable on $[a, b]$.

2. **(M)** The matrix $M(a, b, R)$ is invertible.

Now we state our main result for the estimator $\hat{\alpha}^{(W)}$. Throughout, $\overset{D}{\rightarrow}$ denotes convergence in distribution, and limiting and order relations are always meant as $n \to \infty$ if not specified otherwise.

**Theorem 1.** Assume that the conditions $Q_1 - Q_3$ are satisfied for the underlying distribution and suppose that the quantile function $Q$ admits the representation (3). Moreover, assume the conditions (R) and (M), and assume also that the percentiles $s_j$ are chosen from a closed set $U = [a, b]$, $0 < a < b < 1$, such that $s_j = j/n$, $j = \lceil na \rceil, \ldots, \lceil nb \rceil$, and $\tilde{p} > p$. Then

$$\sqrt{n}(\hat{\alpha}^{(W)}_n - \alpha) \overset{D}{\rightarrow} N(0, V),$$

(4)
where

\[
V = \int_a^b \int_a^b \frac{G_R(s)G_R(t)((1-s) \wedge (1-t) - (1-s)(1-t))}{Q(1-s)Q(1-t)fQ(1-s)fQ(1-t)} ds dt.
\] (5)

The proof is in Chapter 3.

2. Simulation results

In order to make a comparison with existing proposals, simulations were done performed by the Matlab software. The samples were generated from the strict Pareto model \( L \equiv 1 \) in (1) and from the Hall model (2). The Hill, Pickands, DEdH (Dekkers, Einmahl and de Haan) and the weighted least squares (WLS) estimators were included in the simulation study. We used the values \( n = 5000, a = 0.001, b = 0.4 \) and \( \tilde{\rho} = 1, 2, 3 \), and the weight function \( R(s) = s/500 \) for the WLS estimator. In case of \( R \equiv 1 \), we refer to as ordinary least squares (OLS) estimator. The tail indexes were chosen between 0.5 and 20. For the Hill, Pickands and DEdH estimators the simulations were done for sample size \( n = 5000 \) and sample fraction size \( k_n = 200 \). All the simulations were repeated 1000 times.

Tables 1 and 2 contains the empirical mean square errors (MSE) and the average simulated estimates (mean) for the strict Pareto model. We conclude that in the submodel \( L \equiv 1 \) for all \( \alpha \) values, the WLS estimator performs better than the other estimators investigated.

Tables 3 and 4 presents the simulation results for the Hall model. Specifically, we used the parameters \( D_1 = 0.4, D_2 = 1 \) and \( \beta = 0.01 \). We see from Table 3 that the WLS estimator performs better than the other estimators, and the OLS estimator is competitive with the Hill estimator especially for \( \tilde{\rho} = 3 \).
Table 1: Empirical mean square errors (MSE) of tail index estimates for the Pareto model and for sample size \( n = 5000 \).

| \( \alpha \) | \( p = 1 \) | \( p = 2 \) | \( p = 3 \) | \( p = 1 \) | \( p = 2 \) | \( p = 3 \) | \( \text{Hill} \) | \( \text{Pickands} \) | DEdh |
|---|---|---|---|---|---|---|---|---|---|
| 0.5 | 0.00049 | 0.00068 | 0.00094 | 0.00065 | 0.00099 | 0.00135 | 0.00117 | 0.01788 | 0.00678 |
| 0.8 | 0.00183 | 0.00172 | 0.00226 | 0.00161 | 0.00236 | 0.00325 | 0.00335 | 0.02146 | 0.00833 |
| 1 | 0.00176 | 0.00235 | 0.00360 | 0.00242 | 0.00360 | 0.00520 | 0.00545 | 0.02408 | 0.01068 |
| 1.2 | 0.00282 | 0.00393 | 0.00529 | 0.00364 | 0.00536 | 0.00736 | 0.00753 | 0.02512 | 0.01219 |
| 1.5 | 0.00451 | 0.00612 | 0.00857 | 0.00526 | 0.00867 | 0.01188 | 0.01052 | 0.03013 | 0.01608 |
| 1.8 | 0.00604 | 0.00793 | 0.01139 | 0.00769 | 0.01118 | 0.01533 | 0.01680 | 0.03549 | 0.02169 |
| 2 | 0.00763 | 0.01049 | 0.01492 | 0.01084 | 0.01605 | 0.02290 | 0.02019 | 0.03498 | 0.02541 |
| 3 | 0.01766 | 0.02420 | 0.03458 | 0.02353 | 0.03498 | 0.04793 | 0.04665 | 0.06398 | 0.04971 |
| 4 | 0.02913 | 0.04072 | 0.05895 | 0.03926 | 0.05864 | 0.08059 | 0.0807 | 0.09434 | 0.08906 |
| 5 | 0.04768 | 0.06347 | 0.09654 | 0.06409 | 0.09496 | 0.13097 | 0.11472 | 0.13557 | 0.12162 |
| 5.5 | 0.05504 | 0.07689 | 0.10653 | 0.07406 | 0.11049 | 0.15147 | 0.14250 | 0.16283 | 0.14423 |
| 6 | 0.07169 | 0.10385 | 0.14146 | 0.08924 | 0.12963 | 0.17102 | 0.17319 | 0.18811 | 0.17577 |
| 10 | 0.19117 | 0.26276 | 0.37525 | 0.23346 | 0.33953 | 0.45505 | 0.52518 | 0.55818 | 0.52767 |
| 15 | 0.40251 | 0.53582 | 0.80273 | 0.58201 | 0.88450 | 1.22679 | 1.16978 | 1.16751 | 1.17691 |
| 20 | 0.79263 | 1.09561 | 1.57934 | 0.99691 | 1.43447 | 1.91671 | 2.10075 | 1.98171 | 2.10166 |

Table 2: Average simulated tail index estimates (Mean) for sample size \( n = 5000 \) and for the Pareto model.

| \( \alpha \) | \( p = 1 \) | \( p = 2 \) | \( p = 3 \) | \( \text{Hill} \) | \( \text{Pickands} \) | DEdh |
|---|---|---|---|---|---|---|
| 0.5 | 0.50996 | 0.50123 | 0.50257 | 0.50304 | 0.50402 | 0.50508 | 0.50417 | 0.49542 | 0.48967 |
| 0.8 | 0.80139 | 0.80252 | 0.80366 | 0.80557 | 0.80729 | 0.80902 | 0.80028 | 0.80177 | 0.78368 |
| 1 | 1.00148 | 1.00163 | 1.00246 | 1.00531 | 1.00711 | 1.00910 | 1.00182 | 1.00478 | 0.98694 |
| 1.2 | 1.20160 | 1.20180 | 1.20256 | 1.20661 | 1.20894 | 1.21149 | 1.19791 | 1.19525 | 1.18559 |
| 1.5 | 1.50234 | 1.50246 | 1.50263 | 1.50916 | 1.51232 | 1.51584 | 1.50177 | 1.49290 | 1.48542 |
| 1.8 | 1.80561 | 1.80738 | 1.80828 | 1.81250 | 1.81581 | 1.81866 | 1.80135 | 1.80158 | 1.78726 |
| 2 | 2.00607 | 2.00849 | 2.01274 | 2.01694 | 2.02076 | 2.02678 | 2.00450 | 2.00439 | 1.98855 |
| 3 | 3.00475 | 3.00857 | 3.00709 | 3.01346 | 3.01745 | 3.02289 | 3.00717 | 3.00253 | 2.99007 |
| 4 | 4.00635 | 4.00994 | 4.01748 | 4.02856 | 4.03903 | 4.04968 | 3.98594 | 3.98685 | 3.96631 |
| 5 | 5.00734 | 5.01072 | 5.01176 | 5.02099 | 5.02734 | 5.03462 | 5.00493 | 5.01250 | 4.98503 |
| 5.5 | 5.32163 | 5.52344 | 5.55638 | 5.54912 | 5.56201 | 5.57619 | 5.49843 | 5.49632 | 5.48765 |
| 6 | 6.01070 | 6.32096 | 6.30380 | 6.02542 | 6.05761 | 6.07267 | 6.00263 | 6.01285 | 5.98713 |
| 10 | 10.03551 | 10.0453 | 10.04212 | 10.06879 | 10.0851 | 10.0999 | 9.99713 | 10.0411 | 9.98123 |
| 15 | 15.00041 | 15.02029 | 15.05347 | 15.07633 | 15.11221 | 15.14596 | 15.05984 | 15.02914 | 15.0449 |
| 20 | 20.0481 | 20.05749 | 20.09294 | 20.11033 | 20.14008 | 20.17114 | 20.01204 | 20.04928 | 19.99807 |
3. Proof of Theorem 1

Let $q_n(s)$ be the quantile process defined as

$$q_n(s) = \sqrt{n}(Q_n(s) - Q(s)), \quad 0 < s < 1.$$
The proof is based on the strong approximation of the quantile process.

**Theorem 2.** (Csórgő and Révész (1978), Theorem 6.) Suppose that the conditions $Q_1$ and $Q_2$ are satisfied. Then on some probability space one can define a sequence \( \{B_n(t) : 0 \leq t \leq 1\}_{n=1}^{\infty} \) of Brownian bridges such that

\[
\sup_{\delta_n \leq s \leq 1 - \delta_n} |fQ(s)q_n(s) - B_n(s)| \overset{a.s.}{\sim} O(n^{-1/2} \log n),
\]

where \( \delta_n = 25n^{-1} \log n \).

**Proof of Theorem 2.** We assume that the random variables \( X_1, X_2, \ldots \) are defined on the probability space given in Theorem 2. By a simple calculation,

\[
X'WX = \begin{bmatrix}
\sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} & \sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} & 2 \sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} \\
\sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} & \sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} & 2 \sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} \\
\vdots & \vdots & \vdots \\
\sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} & \sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor} & 2 \sum_{j=\lfloor na \rfloor}^{\lfloor nb \rfloor}
\end{bmatrix}
\]

By Riemann sum approximation, we get

\[
\lim_{n \to \infty} n^{-1}X'WX = M(a, b, R)
\]

\[
:= \begin{bmatrix}
\int_a^b & \int_a^b & -2 \int_a^b \\
\int_a^b & \int_a^b & -2 \int_a^b \\
\vdots & \vdots & \vdots \\
\int_a^b & \int_a^b & -2 \int_a^b
\end{bmatrix}
\]

Set \( \varepsilon := (\varepsilon(s_{\lfloor na \rfloor}), \ldots, \varepsilon(s_{\lfloor nb \rfloor}))' \) and

\[
y' := (\log Q(1 - s_{\lfloor na \rfloor}), \ldots, \log Q(1 - s_{\lfloor nb \rfloor})).
\]

Then we have \( y^* = X\beta_\tilde{p}, \beta_\tilde{p} = (X'WX)^{-1}X'Wy^* \) and hence \( \alpha = \epsilon_1' \beta_\tilde{p} = \epsilon_1'(X'WX)^{-1}X'Wy^* \). It follows that \( \varepsilon = y - y^* \) and

\[
\sqrt{n}(\hat{\alpha}_n^{(W)} - \alpha) = \frac{1}{\sqrt{n}} \epsilon_1'(n^{-1}X'WX)^{-1}X'W\varepsilon = Y_n + A_n,
\]

where \( Y_n = n^{-1/2} \epsilon_1'M(a, b, R)^{-1}X'W\varepsilon \) and

\[A_n = n^{-1/2} \epsilon_1'(n^{-1}X'WX)^{-1} - M(a, b, R)^{-1})X'W\varepsilon.\]
A straightforward calculation yields

\[ Y_n = \frac{1}{\sqrt{n}} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} \varepsilon(s_j)G_R(s_j). \]  

(7)

The main point of the proof is to show that

\[ \frac{1}{\sqrt{n}} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} \varepsilon(s_j)G_R(s_j) \overset{D}{\rightarrow} N(0, V). \]  

(8)

With \( \gamma_n(s) := (Q_n(1 - s) - Q(1 - s))/Q(1 - s) \), the residual process can be written as \( \varepsilon(s) = \log(1 + \gamma_n(s)) \). Set \( \eta(x) := \log(1 + x) - x \), and let \( C \) and \( \delta \) be some constants such that \( \eta(x) \leq Cx^2 \), if \( |x| \leq \delta \). Then we obtain

\[ Y_n = Y_{n,1} + A_{n,1}, \]

where

\[ Y_{n,1} = \frac{1}{\sqrt{n}} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} \gamma_n(s_j)G_R(s_j), \quad A_{n,1} = \frac{1}{\sqrt{n}} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} \eta(\gamma_n(s_j))G_R(s_j). \]

First we show that \( A_{n,1} = o_P(1) \). On the event

\[ E_n := \left\{ \max_{\lceil na \rceil \leq j \leq \lfloor nb \rfloor} |\gamma_n(s_j)| \leq \delta \right\}, \]

we have

\[ |A_{n,1}| \leq C \sqrt{n} \max_{\lceil na \rceil \leq j \leq \lfloor nb \rfloor} \gamma_n^2(s_j) \frac{1}{n} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} |G_R(s_j)|. \]

With \( \kappa_1 := \sup_{1-b \leq s \leq 1-a} 1/|Q(s)| \), we obtain

\[ \max_{\lceil na \rceil \leq j \leq \lfloor nb \rfloor} \gamma_n^2(s_j) \leq \kappa_1^2 \sup_{1-b \leq s \leq 1-a} (Q_n(s) - Q(s))^2. \]

Set \( e_n(s) := fQ(s)q_n(s) - B_n(s) \). With the Brownian bridges in Theorem 2 and \( \kappa_2 := \sup_{1-b \leq s \leq 1-a} 1/fQ(s) \) we get

\[ \sup_{1-b \leq s \leq 1-a} |Q_n(s) - Q(s)| = \frac{1}{\sqrt{n}} \sup_{1-b \leq s \leq 1-a} \frac{|e_n(s) + B_n(s)|}{fQ(s)} \]

\[ \leq \frac{\kappa_2}{\sqrt{n}} \sup_{1-b \leq s \leq 1-a} (|e_n(s)| + |B_n(s)|). \]
It follows that
\[ \sqrt{n} \max_{[na] \leq j \leq [nb]} \gamma_n^2(s_j) \leq \frac{\kappa_1^2 \kappa_2^2}{\sqrt{n}} \left( \sup_{1-b \leq s \leq 1-a} |e_n(s)| + \sup_{1-b \leq s \leq 1-a} |B_n(s)| \right)^2. \]

Applying Theorem 2, we obtain
\[ \sqrt{n} \max_{[na] \leq j \leq [nb]} \gamma_n^2(s_j) = o_P(1). \] This, in combination with \( P(E_n) \to 0 \) and \( \frac{1}{n} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} |G_R(s_j)| \to \int_a^b |G_R(s)| ds \) implies \( A_{n,1} = o_P(1) \).

Now we decompose \( Y_{n,1} = Y_{n,2} + A_{n,2} \), where
\[
Y_{n,2} = \frac{1}{n} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} \frac{B_n(1-s_j)G_R(s_j)}{fQ(1-s_j)Q(1-s_j)},
\]
\[
A_{n,2} = \frac{1}{n} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} \frac{e_n(1-s_j)}{fQ(1-s_j)Q(1-s_j)}G_R(s_j).
\]

To prove that \( A_{n,2} = o_P(1) \), we use the inequality
\[ A_{n,2} \leq \kappa_3 \sup_{1-b \leq s \leq 1-a} |e_n(s)| \frac{1}{n} \sum_{j=\lceil na \rceil}^{\lfloor nb \rfloor} |G_R(s_j)|, \]
where
\[ \kappa_3 = \sup_{1-b \leq s \leq 1-a} 1/|fQ(s)Q(s)|. \]

By Theorem 2 we have \( A_{n,2} = o_P(1) \). We prove that the limit of \( Y_{n,2} \) is \( N(0, V) \) given in (4). By the distributional equality
\[
Y_{n,2} \xrightarrow{D} \int_a^b \frac{B(1-s)G_R(s)}{fQ(1-s)Q(1-s)} ds, \quad n = 1, 2, \ldots,
\]
where \( B(\cdot) \) is a Brownian bridge process, we obtain
\[
Y_{n,2} \xrightarrow{D} \int_a^b \frac{B(1-s)G_R(s)}{fQ(1-s)Q(1-s)} ds.
\]

The variance of the limit random variable is described in (5).
The last step is to prove that $A_n = o_P(1)$. Let $(v^*_n, v_{0,n}, \ldots, v^*_{\tilde{p},n})$ be the first row of $(n^{-1}X'WX)^{-1} - M(a, b, R)^{-1}$. Using statement (6), we have $(v^*_n, v_{0,n}, \ldots, v^*_{\tilde{p},n}) \to 0$. Set

$$G^{(n)}(u) := R(u)(-v^*_n \log u + v_{0,n} + 2 \sum_{k=1}^{\tilde{p}} v_{k,n} \cos(2\pi ku)), \quad u \in (0, 1).$$

Similarly as in (7),

$$A_n = \frac{1}{\sqrt{n}} \sum_{j=\lfloor na \rceil}^{\lfloor nb \rceil} \varepsilon(s_j) G^{(n)}(s_j)$$

$$= -v^*_n \frac{1}{\sqrt{n}} \sum_{j=\lfloor na \rceil}^{\lfloor nb \rceil} \varepsilon(s_j) R(s_j) \log s_j + v_{0,n} \frac{1}{\sqrt{n}} \sum_{j=\lfloor na \rceil}^{\lfloor nb \rceil} \varepsilon(s_j) R(s_j)$$

$$+ 2 \sum_{k=1}^{\tilde{p}} v_{k,n} \frac{1}{\sqrt{n}} \sum_{j=\lfloor na \rceil}^{\lfloor nb \rceil} \varepsilon(s_j) R(s_j) \cos(2\pi ks_j).$$

Each term in the last sum tends to zero, e.g., in the first term $v^*_n \to 0$ and applying (8), in which $G_R(s_j)$ is replaced by $R(s_j) \log s_j$, the sequence $\frac{1}{\sqrt{n}} \sum_{j=\lfloor na \rceil}^{\lfloor nb \rceil} \varepsilon(s_j) R(s_j) \log s_j$ has a weak limit.

\[ \square \]

**Acknowledgement.** This research was supported by the Ministry of Human Capacities, Hungary grant TUDFO/47138-1/2019-ITM.

**References**

AL-Najafi, A., Viharos, L., Weighted least squares estimators for the Parzen tail index, Available on arXiv: https://arxiv.org/.

Csörgő, M., Révész, P., 1978. Strong approximations of the quantile process. *Ann. Statist.* 6, 882–894.

Dekkers, A. L. M., Einmahl, J. H. J., de Haan, L., 1989. A moment estimator for the index of an extreme-value estimator. *Ann. Statist.* 17, 1833–1855.
Hall, P., 1982. On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc. Ser. B* **44**, 37–42.

Hill, B. M., 1975. A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163–1174.

Holan, S. H., McElroy, T. S., 2010. Tail exponent estimation via broadband log density-quantile regression. *J. Statist. Plann. Inference* **140**, 3693–3708.

Parzen, E., 2004. Quantile probability and statistical data modeling. *Statist. Sci.* **19**, 652–662.

Pickands III, J., 1975. Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119–131.