Estimation of a tail index based on minimum density power divergence

Moosup Kim, Sangyeol Lee*

Department of Statistics, Seoul National University, Seoul, 151-742, Republic of Korea

Received 25 January 2007
Available online 21 February 2008

Abstract

In this paper, we consider the minimum density power divergence estimator for the tail index of heavy tailed distributions in strong mixing processes. It is shown that the estimator is consistent and asymptotically normal under regularity conditions. The simulation results demonstrate that the estimator is robust in the presence of outliers.

AMS subject classifications: 62M10; 62F12

Keywords: Tail index; Minimum density power divergence estimator; Robustness; Outliers; Strong mixing processes

1. Introduction

The divergence measures are indices used for measuring similarity of or discrepancy between two density functions. The divergence method in the statistical analysis has a very long history, and various divergence measures have been proposed by a number of researchers. Among those, the Kullback–Leibler distance is the most well known, and the \( \phi \)-divergence has long been popular among statisticians: see [7,6,16], and the references therein. Recently, for developing a robust estimation procedure, Basu et al. [2] introduced the density power divergence between two densities \( f \) and \( g \):

\[
d_{\alpha}(g, f) = \begin{cases} 
\int \left\{ f^{1+\alpha}(z) - \left(1 + \frac{1}{\alpha}\right) g(z) f^{\alpha}(z) + \frac{1}{\alpha} g^{1+\alpha}(z) \right\} dz & \text{for } \alpha > 0 \\
\int \log \frac{g(z)}{f(z)} g(z) dz & \text{for } \alpha = 0
\end{cases}
\]

* Corresponding author.
E-mail address: sylee@stats.snu.ac.kr (S. Lee).

© 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2008.02.031
which may be viewed as a particular case of the Bregman divergence (cf. [1]), and designed the minimum density power divergence estimator (MDPDE). The parameter \( \alpha \) controls the trade-off between the efficiency and robustness of the MDPDE: the MDPDE becomes more efficient but less robust against outliers as \( \alpha \) gets closer to zero, while it becomes more robust but less efficient as \( \alpha \) increases. In particular, the \( L_2 \) distance is obtained for \( \alpha = 1 \). An advantage of using the MDPDE over using the minimum Hellinger distance estimator (cf. [3,21]) is that the former can avoid the difficulties, like the problem of an optimal bandwidth choice, which necessarily follow in dealing with the latter. See [2] for more details.

Although the MDPDE was originally studied for an i.i.d. sample, one can easily extend the same estimation procedure to stationary processes. Let \( \{x_i\} \) be a stationary sequence with a common marginal density \( g \) that satisfies

\[
G_n(y) := \frac{1}{n} \sum_{i=1}^{n} I(x_i \leq y) \xrightarrow{p} \int_{-\infty}^{y} g(u)du \quad \text{as} \quad n \to \infty
\]

for each \( y \in \mathbb{R} \), and

\[
\int_{-\infty}^{\infty} h(u)dG_n(u) = \frac{1}{n} \sum_{i=1}^{n} h(x_i) \xrightarrow{P} \int_{-\infty}^{\infty} h(u)g(u)du
\]

for all bounded Borel functions \( h \). Further, let \( \{f_t : t \in \Theta\} \) be a class of densities that is fitted to observations. The MDPDE is defined as the point \( t \) that minimizes the empirical density power divergence:

\[
H_n(t) = \int_{-\infty}^{\infty} f_t^{1+\alpha}(z)dz - \left(1 + \frac{1}{\alpha}\right) \frac{1}{n} \sum_{i=1}^{n} f_t^{\alpha}(x_i).
\]

(1.1)

In an actual derivation, the MDPDE is obtained by solving the equation

\[
U_n(t) = \int_{-\infty}^{\infty} u_t(z)f_t^{1+\alpha}(z)dz - \frac{1}{n} \sum_{i=1}^{n} u_t(x_i)f_t^{\alpha}(x_i) = 0,
\]

(1.2)

where \( u_t(z) = \frac{\partial}{\partial \alpha} \log f_t(z) \) is the score function. It is well known that \( f_t^{\alpha}(x_i) \) in (1.2) functions to make the MDPDE robust against outliers, and the MDPDE is a consistent estimator for \( t^0 := \arg \min \alpha d_\alpha(g, f_t) \) under certain regularity conditions (cf. [2]). For a reference concerning the MDPDE in time series models, see [17].

In this paper, we attempt to apply the above mentioned MDPDE procedure to estimating the tail index of heavy tailed distributions, which belong to the domain of attraction of Fréchet distributions, since conventional estimators like the Hill’s estimator (cf. [12]) are much influenced by abnormal observations: see, for instance, [22]. Later, it will be seen that the minimum density power divergence method not only includes the approach of Vandewalle et al. but also produces more efficient estimates.

In what follows, we describe the procedure used to obtain the MDPDE for the tail index. Let \( \{X_i\} \) be a nonnegative stationary sequence following the distribution \( F \). Since it is well known that a distribution is in the domain of attraction of a Fréchet distribution if and only if the distribution has a regularly varying tail (cf. Theorem 1.6.2 of [14]), we can assume that \( \tilde{F} = 1-F \) is regularly varying at \( \infty \) with the exponent \( -\gamma \) (\( \gamma \) is called the tail index of distribution \( F \)), i.e.,
\[
\lim_{x \to \infty} \frac{\bar{F}(\lambda x)}{\bar{F}(x)} = \lambda^{-\frac{x}{\gamma}} \quad \text{for every } \lambda > 0,
\]

(1.3)

or equivalently, \(\bar{F}(x) = x^{-\frac{1}{\gamma}} l(x)\), where \(l(x)\) is slowly varying at \(\infty\), namely,

\[
\lim_{x \to \infty} \frac{l(\lambda x)}{l(x)} = 1 \quad \text{for all } \lambda > 0.
\]

(1.4)

The idea of constructing the MDPDE for the tail index is to view the logarithms of relative excesses over a given high threshold as exponentially distributed r.v.’s, i.e., we assume that \([X_i]\) satisfies

\[
G_n(x) := \frac{1}{k} \sum_{i=1}^{n} I(\log X_i - \log b(n/k) > x) \overset{P}{\longrightarrow} \int_{x}^{\infty} \gamma^{-1} e^{-u/\gamma} du = e^{-x/\gamma}
\]

for each \(x > 0\), where \(b(x) = F^{-1}(1 - x^{-1})\) \((F^{-1}(x) := \inf\{u : F(u) \geq x\})\), and \(\{k := k_n\}\) is a sequence of positive integers satisfying

\[
k \to \infty \quad \text{and} \quad k = o(n) \quad \text{as } n \to \infty.
\]

Under this assumption, we have

\[
\int_{0}^{\infty} h(u) dG_n(u) \overset{P}{\longrightarrow} \int_{0}^{\infty} h(u) \gamma^{-1} e^{-u/\gamma} du
\]

for all bounded and continuous \(h\) defined on \(\mathbb{R}^+\). Thus, by fitting the exponential model to the logarithms of relative excesses over \(b(n/k)\), we can reformulate the estimating equation in (1.2) as follows:

\[
\int_{0}^{\infty} \left( \frac{z}{t^2} - \frac{1}{t} \right) \left\{ \frac{1}{t} e^{-z/t} \right\}^{1+\alpha} dG_n(z) = \frac{\alpha}{(1 + \alpha)^2 t^{1+\alpha}} - \frac{1}{k} \sum_{i=1}^{n} \left( \frac{Y_i}{t^2} - \frac{1}{t} \right) \frac{1}{t^\alpha} \exp \left( -\frac{\alpha Y_i}{t} \right) I(Y_i > 0) = 0,
\]

where \(Y_i = \log X_i - \log b(n/k)\), and \(\frac{1}{t} e^{-z/t} I(z > 0)\) and \((\frac{z}{t^2} - \frac{1}{t}) I(z > 0)\) are the density of the exponential distribution and its score function, respectively. Since \(b(n/k)\) is unknown, we replace it by \(X_{(k+1)}\), the \((k + 1)\)-th largest value of \(X_1, \ldots, X_n\), and define the MDPDE as the solution of the equation

\[
U_n(t) = -\frac{\alpha}{(1 + \alpha)^2 t^{1+\alpha}} - \frac{1}{k} \sum_{i=1}^{n} \left( \frac{\tilde{Y}_i}{t^2} - \frac{1}{t} \right) \frac{1}{t^\alpha} \exp \left( -\frac{\alpha \tilde{Y}_i}{t} \right) I(\tilde{Y}_i > 0) = 0,
\]

(1.5)

where \(\tilde{Y}_i = \log X_i - \log X_{(k+1)}\). If multiple solutions of (1.5) exist, we choose the one that minimizes the corresponding \(H_n\). It is noteworthy that when \(\alpha = 0\), the estimating equation produces the Hill’s estimator.

The organization of this paper is as follows. In Section 2, we verify that the MDPDE obtained based on (1.5) is consistent and asymptotically normal for a class of strong mixing processes. In Section 3, we provide a simulation result that demonstrates the validity of the MDPDE. All the proofs are provided in Section 4.
2. Asymptotic properties of the MDPDE

In this section, we verify that the MDPDE is consistent and asymptotically normal under some regularity conditions. In what follows, we assume that all r.v.’s are defined on the probability space \((\Omega, \mathcal{F}, P)\), and \([k = k_n]\) denotes a sequence of positive integers such that \(k \to \infty\) and \(k = o(n)\) as \(n \to \infty\). In this study, we only consider the case of \(\alpha > 0\) since the case of \(\alpha = 0\) is already handled by Hsing [13]. The proofs of the results in this section are provided in Section 4.

2.1. Consistency

**Theorem 1.** Suppose that

\[
\frac{1}{k} \sum_{i=1}^{n} \{I(\log X_i > \log b(n/k) + x) - P(\log X_i > \log b(n/k) + x)\} \to^p 0 \tag{2.6}
\]

for every \(x \in \mathbb{R}\). Then, there exists a sequence \(\{\hat{\gamma}_n\}\) satisfying \(U_n(\hat{\gamma}_n) = 0\) such that \(\hat{\gamma}_n \to^p \gamma\).

It is well known that a broad class of strong mixing processes satisfy the conditions in (2.6). In order to obtain the consistency result, stronger conditions are needed as addressed below.

**Corollary 1.** Let

\[
\beta(l) := \sup_{m \in \mathbb{N}} \sup_{A \in \mathcal{F}_m^l, B \in \mathcal{F}_m^\infty} \{|P(A \cap B) - P(A) P(B)|\}, \tag{2.7}
\]

where \(\mathcal{F}_m^l = \sigma\{X_i, X_{i+1}, \ldots, X_m\} (l \leq m)\) and \(\mathcal{F}_m^\infty = \sigma\{X_m, X_{m+1}, \ldots\}\). Suppose that (1.3) holds and there exists a sequence \(\{r_n\}\) of positive integers such that \(r_n = o(k)\) and

\[
\lim_{n \to \infty} m_n \beta(r_n) = 0, \tag{2.8}
\]

where \(m_n = \lfloor n/r_n \rfloor\). Then, (2.6) holds, and subsequently, there exists a sequence \(\{\hat{\gamma}_n\}\) satisfying \(U_n(\hat{\gamma}_n) = 0\) such that \(\hat{\gamma}_n \to^p \gamma\).

2.2. Asymptotic normality

For a bivariate function \(h(x, t)\), we define \(\hat{h}(x, t) = \frac{\partial}{\partial t} h(x, t)\) and \(\tilde{h}(x, t) = \frac{\partial^2}{\partial t^2} h(x, t)\). For notational convenience in handling the argument in (1.5), we introduce

\[
\phi(x, t; \alpha) := \left(\frac{x}{t} - \frac{1}{t}\right) \frac{1}{t^\alpha} \exp\left(-\frac{\alpha x}{t}\right), \quad \psi(x, t; \alpha) := \frac{\alpha}{(1 + \alpha)^{1+\alpha}} - \phi(x, t; \alpha),
\]

and

\[
J(t; \alpha) := \int_0^\infty \psi(x, t; \alpha) \frac{1}{t} e^{-\xi/t} dx = \frac{1 + \alpha^2}{t^{2+\alpha} (1 + \alpha)^3} > 0.
\]

We set

\[
Y_i := \log X_i - \log b(n/k), \quad Y_i^{(\zeta)} := I(\sqrt{k} Y_i > \zeta) \quad \text{for } \zeta \in \mathbb{R}.
\]

Further, we set \(\phi(x) := \phi(x, \gamma; \alpha)\) and denote by \(RV_\eta\) the class of all functions regularly varying at \(\infty\) with the exponent \(\eta\).

In order to achieve the asymptotic normality of the MDPDE, we impose the following regularity conditions on (1.4):
There exist \( \kappa(x) = K \int_1^x t^{r-1} dt \) (\( K \) is finite) and a positive measurable \( g \in RV_\tau (\tau \leq 0) \) such that for all \( \lambda > 1 \),

\[
\lim_{x \to \infty} \frac{l(\lambda x)}{l(x)} - 1 = \kappa(\lambda).
\] (2.9)

Further, \( \sqrt{k} g(b(n/k)) \to 0 \) as \( n \to \infty \).

Furthermore, we assume that there is a sequence of positive integers \( \{r_n\} \) such that

\[
r_n^2 = o(n) \quad \text{and} \quad \lim_{n \to \infty} m_n \beta(\lfloor r_n \rfloor) = 0 \quad \text{for any} \quad 0 < \epsilon < 1,
\] (2.10)

where \( m_n := \lfloor n/r_n \rfloor \).

Note that the condition in (2.9) is a slight modification of Condition (SR2) of [13]. The following is the main result of this subsection.

**Theorem 2.** Suppose that (1.3) and A hold. Then, there exists a sequence of positive integers \( \{r_n\} (r_n \to \infty) \) satisfying (2.10), and

\[
\frac{1}{k} \sum_{i=1}^{n} \left\{ I \left( \log X_i > \log b \left( \frac{n}{\rho k} \right) + x \right) - P \left( \log X_i > \log b \left( \frac{n}{\rho k} \right) + x \right) \right\} \to 0
\] (2.11)

for each \( x \in \mathbb{R} \) and \( \rho \) in some neighborhood \( I \) of 1. Further, suppose that there exist constants \( \chi, \vartheta, \) and \( \omega \) such that as \( n \to \infty \),

\[
\frac{2n}{[\varepsilon r_n] k} \sum_{1 \leq i < j \leq \lfloor \varepsilon r_n \rfloor} \text{Cov}(\phi(Y_i)Y_i^{(0)}, \phi(Y_j)Y_j^{(0)}) \to \chi,
\] (2.12)

\[
\frac{\gamma n}{[\varepsilon r_n] k} \sum_{1 \leq i < j \leq \lfloor \varepsilon r_n \rfloor} \{\text{Cov}(\phi(Y_i)Y_i^{(0)}, Y_j^{(\xi)}) + \text{Cov}(Y_i^{(\xi)}, \phi(Y_j)Y_j^{(0)})\} \to \vartheta,
\] (2.13)

\[
\frac{n \gamma^2}{[\varepsilon r_n] k} \sum_{1 \leq i < j \leq \lfloor \varepsilon r_n \rfloor} \{\text{Cov}(Y_i^{(0)}, Y_j^{(\xi)}) + \text{Cov}(Y_i^{(\xi)}, Y_j^{(0)})\} \to \omega,
\] (2.14)

and

\[
\frac{2n \gamma^2}{[\varepsilon r_n] k} \sum_{1 \leq i < j \leq \lfloor \varepsilon r_n \rfloor} \text{Cov}(Y_i^{(\xi)}, Y_j^{(\xi)}) \to \omega
\] (2.15)

for any \( 0 < \epsilon < 1 \) and \( \xi \in \mathbb{R} \). Then, if for any \( \delta > 0, \epsilon > 0, c_1, c_2, c_3, \) and \( \xi \in \mathbb{R} \),

\[
\frac{m_n}{k} \mathbb{E} W_n^2 I (|W_n| > \delta \sqrt{k}) \to 0 \quad \text{as} \quad n \to \infty,
\] (2.16)

where

\[
W_n = \sum_{j=1}^{\lfloor \varepsilon r_n \rfloor} \left\{ c_1 \left( \phi(Y_j)Y_j^{(0)} - \mathbb{E} \phi(Y_j)Y_j^{(0)} \right) + c_2 \gamma \left( Y_j^{(0)} - \mathbb{E} Y_j^{(0)} \right) \\
+ c_3 \gamma \left( Y_j^{(\xi)} - \mathbb{E} Y_j^{(\xi)} \right) \right\}
\]

it holds that

\[
\sqrt{k} (\hat{\gamma}_n - \gamma) \Rightarrow N(0, V),
\]
where \( V = V_0 + V_1 \),

\[
V_0 = \frac{(1 + \alpha)^2}{(1 + \alpha^2)^2} \left\{ \frac{(1 + \alpha)^4(1 + 4\alpha^2)}{(1 + 2\alpha^3)} - \alpha^2 \right\} \gamma^2,
\]

and

\[
V_1 = \frac{(1 + \alpha)^6\gamma^{4+2\alpha}}{(1 + \alpha^2)^2} \left\{ \chi + \frac{2\alpha}{\gamma^{2+\alpha}(1 + \alpha^2)} + \frac{\alpha^2\omega}{\gamma^{4+2\alpha}(1 + \alpha)^4} \right\}.
\]

**Remark 1.** If \( \frac{r_n^2}{n} = o(k) \), the above mentioned conditions can be slightly relaxed (see Corollary 2). A typical example of the stationary process satisfying all those conditions is the infinite order moving average process with mixing order geometrically decaying to 0 (cf. [10, 19]), which includes ARMA processes.

**Remark 2 (Bias and Asymptotic MSE of \( \hat{y}_n \)).** Suppose that observations are i.i.d. and

\[
l(x) = C \left( 1 + DX^T + O\left( x^n \right) \right), \quad C > 0, D \neq 0, \eta < \tau < 0
\]
as \( x \to \infty \). In this case, \( A \) holds with \( g(x) = \tau Dx^T \) and \( K = 1 \) (except that \( \sqrt{k}g(b(n/k)) \to 0 \) as \( n \to \infty \)). For the details regarding the condition in (2.17), we refer the reader to [8]. By a slight modification of the proofs of Lemma 2 and Theorems 2 and 3, and by using the fact that \( b(x) = x^\gamma C^\psi (1 + O\left( x^\tau \right)) \) (as \( x \to \infty \)), we can write

\[
\hat{y}_n = \gamma + k^{-1/2} \left\{ N_n + \sqrt{k} \tau J^{-1}(\gamma; \alpha) M(\gamma, \alpha, \tau) DC^{-\rho}(k/n)^\rho + o_P(1) \right\}
\]

\[
= \gamma + k^{-1/2} N_n + \tau J^{-1}(\gamma; \alpha) M(\gamma, \alpha, \tau) DC^{-\rho}(k/n)^\rho + o_P \left( \frac{1}{\sqrt{k}} \right),
\]

where \( M(\gamma, \alpha, \tau) = \int_0^\infty e^{-\gamma} \tau \kappa((e^\gamma)^k) d\phi(\gamma) = \frac{1}{\sqrt{\tau^{2+\alpha}} \left( \frac{a}{1+\alpha} - \frac{(\rho+1)(\alpha+\rho)}{(1+\alpha+\rho)} \right), \rho = -\tau \gamma, \) and

\[
N_n \Rightarrow \mathcal{N} \left( 0, \frac{(1 + \alpha)^2}{(1 + \alpha^2)^2} \left\{ \frac{(1 + \alpha)^4(1 + 4\alpha^2)}{(1 + 2\alpha^3)} - \alpha^2 \right\} \gamma^2 \right).
\]

Therefore, the bias and asymptotic MSE of \( \hat{y}_n \) are obtained as

\[
\frac{(1 + \alpha)^3}{1 + \alpha^2} \left\{ \frac{\alpha}{1 + \alpha} - \frac{(\rho + 1)(\alpha + \rho)}{(1 + \alpha + \rho)} \right\} DC^{-\rho}(k/n)^\rho
\]

and

\[
\frac{(1 + \alpha)^6}{(1 + \alpha^2)^2} \left\{ \frac{\alpha}{1 + \alpha} - \frac{(\rho + 1)(\alpha + \rho)}{(1 + \alpha + \rho)} \right\}^2 D^2 C^{-2\rho}(k/n)^{2\rho}
\]

\[
+ \frac{(1 + \alpha)^2}{(1 + \alpha^2)^2} \left\{ \frac{(1 + \alpha)^4(1 + 4\alpha^2)}{(1 + 2\alpha^3)} - \alpha^2 \right\} \gamma^2 \frac{1}{k},
\]

respectively. We can see that the bias of \( \hat{y}_n \) has the same decaying rate as the conventional Hill’s estimator (cf. [11]), and the asymptotic MSE of \( \hat{y}_n \) is greater than that of the Hill’s estimator.

**Remark 3 (Optimal Level k).** The optimal \( k^{\text{opt}} \) is determined as the \( k \) that minimizes the asymptotic MSE in Remark 2. In fact, it can be seen that the rate of \( k^{\text{opt}} \) is \( n^{2\rho/(2\rho+1)} \), which is
identical to that of the Hill’s estimator (cf. [11]). Then, if we set \( k = \lambda n^{2\rho/(2\rho+1)} \), the asymptotic MSE is obtained as \( n^{-2\rho/(2\rho+1)} \) multiplied by the number:

\[
\left(\frac{1+\alpha}{(1+\alpha^2)^2}\right)^6 \left\{ \frac{\alpha}{1+\alpha} - \frac{(\rho+1)(\alpha + \rho)}{(1+\alpha + \rho)} \right\}^2 D^2 C^{-2\rho} \lambda^{2\rho} + \left(\frac{1+\alpha}{(1+\alpha^2)^2}\right)^2 \left\{ \frac{(1+\alpha)^4(1+4\alpha^2)}{(1+2\alpha)^3} - \alpha^2 \right\} \gamma^2 \lambda.
\]

This indicates that \( k_{\text{opt}} = \lambda^* n^{2\rho/(2\rho+1)} \), where \( \lambda^* \) is the \( \lambda \) that minimizes (2.18).

3. Simulation study

In this section, we evaluate the performance of the MDPDE of the tail index through a simulation study. First, we investigate the performance of the MDPDE in the case where observations do not include outliers. To achieve this, we consider the following four cases:

(1) Let \( X_i, i = 1, 2, \ldots, n \), be i.i.d. observations following a \( t \)-distribution with degrees of freedom 2 (its tail index is 0.5).

(2) Let \( X_i = \xi_i + 0.5 \xi_{i-1}, i = 1, 2, \ldots, n \), be a first-order moving average sequence, where \( \xi_i \) are i.i.d. r.v.’s following a \( t \)-distribution with degrees of freedom 2: its tail index is 0.5, and the distribution of \( X_1 \) also has a regularly varying tail with tail index 0.5.

(3) Let \( X_i, i = 1, 2, \ldots, n \), be i.i.d. observations following the Burr distribution:

\[
\bar{F}(x) = \left( \frac{\beta}{\beta + x^{-\tau}} \right)^\lambda
\]

with \( \beta = 1, \lambda = 1, \) and \( \tau = -2 \) (its tail index is 0.5).

(4) We consider the same situation as in Case (3) with \( \lambda = 2 \) and \( \tau = -1 \) (its tail index is still 0.5).

In the above set-up, we evaluate the performance of the MDPDE, the Hill’s estimator, and the bias-reduced estimator proposed by Feuerverger and Hall [8] by comparing their MSE’s. The bias-reduced estimator is obtained through the least squared approach, which is well known to be robust against departures from classical extreme value approximations. In each simulation, the repetition number is 1000.

Tables 1 and 2 show the MSE’s in Cases (1) and (2), respectively. It can be observed that the MSE’s of MDPDE are greater than those of the other two estimators, and the bias-reduced estimator outperforms the other two estimators. These results appeal to our intuition. Meanwhile, Table 5 shows the relative efficiency of the MDPDE with respect to the Hill estimator theoretically obtained based on Theorem 2, where the relative efficiency is defined as the asymptotic variance of the Hill estimator divided by that of the MDPDE: as anticipated, the MDPDE loses efficiency as \( \alpha \) increases.

In fact, Cases (3) and (4) are considered to see the effect of the second-order regularly varying parameter \( \tau \) on the performance of the three tail index estimators. The results summarized in Tables 3 and 4 demonstrate that all the estimators in Case (3) outperform those in Case (4). This is natural since the \( \tau \) close to 0 damages the step of viewing the logarithms of relative excesses as exponentially distributed in constructing the MDPDE.

We now turn our attention to the case where observations are contaminated by outliers. In order to achieve our aim, we consider the following cases:
Table 1
MSE of estimators for Case (1)

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>Hill Bias-reduced</th>
<th>MDPDE/α</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>1000</td>
<td>50</td>
<td>0.0111</td>
<td>0.0127</td>
</tr>
<tr>
<td>100</td>
<td>0.0316</td>
<td>0.0352</td>
<td>0.0371</td>
</tr>
<tr>
<td>150</td>
<td>0.0857</td>
<td>0.0392</td>
<td>0.1088</td>
</tr>
<tr>
<td>200</td>
<td>0.1934</td>
<td>0.0414</td>
<td>0.2300</td>
</tr>
<tr>
<td>3000</td>
<td>50</td>
<td>0.0099</td>
<td>0.0185</td>
</tr>
<tr>
<td>100</td>
<td>0.0063</td>
<td>0.0085</td>
<td>0.0071</td>
</tr>
<tr>
<td>150</td>
<td>0.0072</td>
<td>0.0076</td>
<td>0.0082</td>
</tr>
<tr>
<td>200</td>
<td>0.0110</td>
<td>0.0125</td>
<td>0.0156</td>
</tr>
<tr>
<td>250</td>
<td>0.0194</td>
<td>0.0137</td>
<td>0.0265</td>
</tr>
<tr>
<td>300</td>
<td>0.0328</td>
<td>0.0130</td>
<td>0.0372</td>
</tr>
</tbody>
</table>

Table 2
MSE of estimators for Case (2)

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>Hill</th>
<th>MDPDE/α</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>1000</td>
<td>50</td>
<td>0.0136</td>
<td>0.0140</td>
</tr>
<tr>
<td>100</td>
<td>0.0120</td>
<td>0.0162</td>
<td>0.0256</td>
</tr>
<tr>
<td>150</td>
<td>0.0517</td>
<td>0.0625</td>
<td>0.0937</td>
</tr>
<tr>
<td>200</td>
<td>0.1360</td>
<td>0.1700</td>
<td>0.2398</td>
</tr>
<tr>
<td>3000</td>
<td>50</td>
<td>0.0123</td>
<td>0.0130</td>
</tr>
<tr>
<td>100</td>
<td>0.0058</td>
<td>0.0056</td>
<td>0.0062</td>
</tr>
<tr>
<td>150</td>
<td>0.0047</td>
<td>0.0045</td>
<td>0.0052</td>
</tr>
<tr>
<td>200</td>
<td>0.0041</td>
<td>0.0050</td>
<td>0.0066</td>
</tr>
<tr>
<td>250</td>
<td>0.0067</td>
<td>0.0091</td>
<td>0.0131</td>
</tr>
<tr>
<td>300</td>
<td>0.0135</td>
<td>0.0171</td>
<td>0.0255</td>
</tr>
</tbody>
</table>

Table 3
MSE of estimators for Case (3)

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>Hill</th>
<th>Bias-reduced</th>
<th>MDPDE/α</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>1000</td>
<td>50</td>
<td>0.0102</td>
<td>0.0285</td>
<td>0.0083</td>
</tr>
<tr>
<td>100</td>
<td>0.0061</td>
<td>0.0147</td>
<td>0.0069</td>
<td>0.0081</td>
</tr>
<tr>
<td>150</td>
<td>0.0079</td>
<td>0.0153</td>
<td>0.0088</td>
<td>0.0121</td>
</tr>
<tr>
<td>200</td>
<td>0.0134</td>
<td>0.0142</td>
<td>0.0169</td>
<td>0.0230</td>
</tr>
<tr>
<td>3000</td>
<td>50</td>
<td>0.0093</td>
<td>0.0158</td>
<td>0.0089</td>
</tr>
<tr>
<td>100</td>
<td>0.0054</td>
<td>0.0072</td>
<td>0.0049</td>
<td>0.0056</td>
</tr>
<tr>
<td>150</td>
<td>0.0034</td>
<td>0.0052</td>
<td>0.0033</td>
<td>0.0044</td>
</tr>
<tr>
<td>200</td>
<td>0.0028</td>
<td>0.0057</td>
<td>0.0029</td>
<td>0.0046</td>
</tr>
<tr>
<td>250</td>
<td>0.0027</td>
<td>0.0060</td>
<td>0.0032</td>
<td>0.0039</td>
</tr>
<tr>
<td>300</td>
<td>0.0029</td>
<td>0.0049</td>
<td>0.0038</td>
<td>0.0043</td>
</tr>
</tbody>
</table>

(5) Let \(Z_i\) and \(U_i, i = 1, \ldots, n\), be independent r.v.'s following a \(t\)-distribution with degrees of freedom 2 and \(\frac{1}{3}\), respectively (the tail index \(\gamma\) is 0.5 and 3, respectively). In this case, we consider the situation in which the \(Z_i\) are contaminated by the \(U_i\), and the MDPDE of \(\gamma\) is obtained from the observations \(X_i = (1 - \delta_i)Z_i + \delta_iU_i\), where \(\delta_i\) are i.i.d. Bernoulli r.v.'s with the success probability 0.05, and \(\{\delta_i\}, \{Z_i\}, \text{ and } \{U_i\}\) are assumed to be all independent.
Tables 4 and 5 exhibit the MSE’s of the Hill estimator, the bias-reduced estimator, and the MDPDE for Cases (5) and (6), respectively. In almost all cases, the MSE of the MDPDE appears to be less than that of the Hill estimator and that of the bias-reduced estimator, which implies that the MDPDE is more robust against outliers than the other estimators. Further, as might be anticipated, we can see that the MSE has a tendency to decrease as $\alpha$ increases, which confirms that the $\alpha$ properly controls the degree of robustness of the MDPDE. Our findings in this simulation study enable us to conclude that the MDPDE is a promising robust estimator for the tail index parameter when the data set is contaminated by outliers.

(6) Let $Z_i = \xi_i + 0.5 \xi_{i-1}$, where $\xi_i$ are i.i.d. r.v.’s following a $t$-distribution with degrees of freedom 2. All the remaining other parts are the same as in Case (5).
Table 7
MSE of estimators for Case (6)

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>Hill</th>
<th>MDPDE/α</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.1 0.3 0.5 0.7 0.9 1</td>
</tr>
<tr>
<td>1000</td>
<td>50</td>
<td>1.1498</td>
<td>0.5624 0.1666 0.1001 0.0947 0.0900 0.1009</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.4625</td>
<td>0.2361 0.1146 0.0925 0.1090 0.1117 0.1214</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.4200</td>
<td>0.2683 0.1900 0.1906 0.2194 0.2302 0.2498</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.4930</td>
<td>0.3932 0.3509 0.4011 0.4478 0.4821 0.5242</td>
</tr>
<tr>
<td>3000</td>
<td>50</td>
<td>5.8982</td>
<td>4.4672 2.2522 1.3905 1.0642 0.9123 0.9103</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>2.1701</td>
<td>1.1928 0.3813 0.2123 0.1773 0.1664 0.1631</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>1.1581</td>
<td>0.5548 0.1783 0.1019 0.0888 0.0878 0.0965</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.8035</td>
<td>0.3602 0.1224 0.0891 0.0796 0.0820 0.0818</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>0.6015</td>
<td>0.2811 0.1072 0.0828 0.0864 0.0894 0.0921</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.5077</td>
<td>0.2446 0.1174 0.0927 0.1014 0.1076 0.1071</td>
</tr>
</tbody>
</table>

4. Proofs

4.1. Proof of the consistency of the MDPDE

The following lemma is useful for proving the consistency of the MDPDE.

**Lemma 1.** Suppose that for each \( x \in \mathbb{R}^+ \),

\[ \tilde{G}_n(x) := \frac{1}{k} \sum_{i=1}^{n} I \{ \tilde{Y}_i > x \} \xrightarrow{P} e^{-\frac{x}{\gamma}}. \]  

(4.19)

Then,

\[ \int_{0}^{\infty} h(z) d\tilde{G}_n(x) = \frac{1}{k} \sum_{i=1}^{n} h(\tilde{Y}_i) I(\tilde{Y}_i > 0) \xrightarrow{P} \int_{0}^{\infty} h(x) \frac{1}{\gamma} e^{-\frac{x}{\gamma}} dx, \]  

(4.20)

for each \( h \) is a continuous and bounded function defined on \( \mathbb{R}^+ \). Further, we have a sequence \( \{\hat{\gamma}_n\} \) satisfying \( U_n(\hat{\gamma}_n) = 0 \) such that \( \hat{\gamma}_n \xrightarrow{P} \gamma \).

**Proof.** Set

\[ \frac{1}{k} \sum_{i=1}^{k} I \{ \tilde{Y}_{(i)} \leq x \} = 1 - \frac{1}{k} \sum_{i=1}^{n} I \{ \tilde{Y}_i > x \} = 1 - \tilde{G}_n(x), \quad (x > 0), \]

where \( \tilde{Y}_{(i)} \) is the \( i \)-th largest value of \( \tilde{Y}_1, \ldots, \tilde{Y}_n \). Let \( \{\tilde{G}_{n'}\} \) be any subsequence of \( \{\tilde{G}_n\} \). Then, using the conventional diagonal method, we can choose a further subsequence \( \{\tilde{G}_{n''}\} \) such that with probability 1,

\[ \lim_{n'' \to \infty} \left| \tilde{G}_{n''}(x) - e^{-\frac{x}{\gamma}} \right| = 0 \quad \text{for every} \, x \in \mathbb{Q}^+, \]  

(4.21)

where \( \mathbb{Q}^+ \) is the set of positive rational numbers. Since \( x \mapsto \tilde{G}_{n''}(x) \) is non-increasing and \( x \mapsto e^{-\frac{x}{\gamma}} \) is continuous in \( x \), with probability 1, the above convergence holds uniformly for \( x \in \mathbb{R}^+ \). Hence, in view of Billingsley [4], we have that for all bounded continuous functions \( h \),

\[ \frac{1}{k''} \sum_{i=1}^{k''} h(\tilde{Y}_{(i)}) = \int_{0}^{\infty} h(z) d\tilde{G}_{n''}(x) \xrightarrow{a.s.} \int_{0}^{\infty} h(x) \frac{1}{\gamma} e^{-\frac{x}{\gamma}} dx. \]
where $k'' = k_n''$. Therefore, (4.20) is asserted.

Now, to prove the existence of a consistent solution of $U_n(t) = 0$, we follow the lines of the proof of Theorem 6.4.1 of [15]. By Taylor’s theorem, we can write

$$H_n(t) - H_n(\gamma) = U_n(\gamma)(t - \gamma) + \frac{\dot{U}_n(\gamma)}{2}(t - \gamma)^2 + \frac{\ddot{U}_n(t^*)}{6}(t - \gamma)^3$$

for some $t^*$ between $t$ and $\gamma$. Further, both $\psi(x, \gamma; \alpha)$ and $\dot{\psi}(x, \gamma; \alpha)$ are bounded and continuous functions in $x \in \mathbb{R}^+$, and $\sup_{x > 0} \sup_{t \in (\gamma - \delta, \gamma + \delta)} |\ddot{\psi}(x, t; \alpha)| < \infty$ for some $\delta \in (0, \gamma)$. By using these facts, (4.20), and the positivity of $J(\gamma; \alpha)$, we can readily verify the theorem. The details are omitted for brevity. □

**Proof of Theorem 1.** From (2.6), we have that for each $x \in \mathbb{R}$,

$$G_n(x) := \frac{1}{k} \sum_{i=1}^{n} I \left( \log X_i - \log b(n/k) > x \right) \xrightarrow{P} e^{-\frac{x}{\gamma}}$$

which in turn implies

$$\log X_{(k+1)} - \log b(n/k) \xrightarrow{P} 0$$

(cf. the proof of Theorem 2.2 in [13]). Since both $x \mapsto G_n(x)$ and $x \mapsto e^{-\frac{x}{\gamma}}$ decrease as $x \to \infty$, and $x \mapsto e^{-\frac{x}{\gamma}}$ is continuous, we have

$$\sup_{x \in I} \left| G_n(x) - e^{-\frac{x}{\gamma}} \right| \xrightarrow{P} 0 \quad \text{as } n \to \infty$$

for every bounded interval $I$. Thus, by replacing $\log b(n/k)$ with $\log X_{(k+1)}$, we have

$$\frac{1}{k} \sum_{i=1}^{n} I \left( \log X_i - \log X_{(k+1)} > x \right) \xrightarrow{P} e^{-\frac{x}{\gamma}},$$

which implies (4.19). This completes the proof. □

**Proof of Corollary 1.** According to Theorem 3.1 of [13], we have (2.6). Hence, by Theorem 1, we assert the corollary. □

### 4.2. Proof of the asymptotic normality of MDPDE

In this subsection, we prove the asymptotic normality of the MDPDE. The idea of the proof heavily depends on that of the Hill estimator (cf. [13]). For a bivariate function $h(x, t)$, we put $h'(x, t) = \frac{\partial}{\partial x} h(x, t)$ and $h''(x, t) = \frac{\partial^2}{\partial x^2} h(x, t)$. Further, we set

$$\nu_1(\gamma; \alpha) := \int_{0}^{\infty} \phi(y) \frac{1}{\gamma} e^{-\frac{y}{\gamma}} \, dy = -\frac{\alpha}{(1 + \alpha)^2 \gamma^{1 + \alpha}}$$

and

$$\nu_2(\gamma; \alpha) := \int_{0}^{\infty} \phi^2(y) \frac{1}{\gamma} e^{-\frac{y}{\gamma}} \, dy = \frac{1 + 4\alpha^2}{\gamma^{2 + 2\alpha}(1 + 2\alpha)^3}.$$  

The following lemma is concerned with the moments of the functions of $Y_i = \log X_i - \log b(n/k)$, which is crucial for verifying the asymptotic normality of the MDPDE.
Lemma 2. Under A, we have
\[
E I(\sqrt{k} Y_i > \zeta) = \frac{k}{n} \left\{ 1 - \frac{\zeta}{\gamma \sqrt{k}} + o \left( \frac{1}{\sqrt{k}} \right) \right\},
\]
\[
E \Phi(Y_i) I(\sqrt{k} Y_i > \zeta) = \frac{k}{n} \{v_1(\gamma; \alpha) + o(1)\},
\]
\[
E \Phi^2(Y_i) I(Y_i > 0) = \frac{k}{n} \{v_2(\gamma; \alpha) + o(1)\}.
\]

and,
\[
E \Phi(Y_i) I(Y_i > 0) = \frac{k}{n} \{v_1(\gamma; \alpha) + o \left( \frac{1}{\sqrt{k}} \right) \}. \tag{4.23}
\]

\textbf{Proof.} We only provide the proof for (4.23) since the remaining part of the statements can be proven similarly. Note that
\[
E \Phi(Y_1) I(Y_1 > 0) = \int_{b(n/k)}^{\infty} \Phi \left( \log \frac{x}{b(n/k)} \right) dF(x)
\]
\[
= \Phi(0) \tilde{F}(b(n/k)) + \int_{b(n/k)}^{\infty} \tilde{F}(x) d\Phi \left( \log \frac{x}{b(n/k)} \right), \tag{4.24}
\]
where the second term in (4.24) is obtained by the integration by parts. Let \( y = x/b(n/k) \). Then the second term in (4.24) is rewritten as follows:
\[
\int_{1}^{\infty} \tilde{F}(b(n/k)y) \phi'(\log y) \frac{dy}{y} = \tilde{F}(b(n/k)) \int_{1}^{\infty} \tilde{F}(b(n/k)y) \frac{\phi'(\log y) dy}{y}
\]
\[
= \tilde{F}(b(n/k)) \int_{1}^{\infty} y^{-(1+\frac{\gamma}{\alpha})} \phi'(\log y) \frac{l(b(n/k)y)}{l(b(n/k))} \, dy. \tag{4.25}
\]

In view of Goldie and Smith \[9\], we can write
\[
\int_{1}^{\infty} y^{-(1+\frac{\gamma}{\alpha})} \phi'(\log y) \frac{l(b(n/k)y)}{l(b(n/k))} \, dy = \int_{1}^{\infty} y^{-(1+\frac{\gamma}{\alpha})} \phi'(\log y) \, dy + Mg(b(n/k)) + o \left( \frac{1}{\sqrt{k}} \right),
\]
where \( M = \int_{0}^{\infty} e^{-\gamma} \kappa(e^y) \, d\phi(y) \). By using the integration by parts, it follows from (4.24) and (4.25) that
\[
E \Phi(Y_1) I(Y_1 > 0) = \tilde{F}(b(n/k)) \left\{ v_1(\gamma; \alpha) + Mg(b(n/k)) + o \left( \frac{1}{\sqrt{k}} \right) \right\}.
\]
Since \( \tilde{F}(b(n/k)) = \frac{k}{n} \{1 + o(g(b(n/k)))\} \) (cf. the arguments up to (3.1) in \[20\]) and \( \sqrt{k}g(b(n/k)) \to 0 \) as \( n \to \infty \),
\[
E \Phi(Y_1) I(Y_1 > 0) = \frac{k}{n} \left\{ v_1(\gamma; \alpha) + o \left( \frac{1}{\sqrt{k}} \right) \right\}.
\]

Hence, the proof is completed. \( \square \)
Here we prove a lemma and a series of theorems to establish the asymptotic normality of the MDPDE.

**Lemma 3.** Suppose that there exists a neighborhood \( I \) of 1 such that for all \( x \in \mathbb{R} \) and \( \rho \) in \( I \),

\[
\frac{1}{k} \sum_{i=1}^{n} \left\{ I \left( \log X_i > \log b \left( \frac{n}{\rho k} \right) + x \right) - P \left( \log X_i > \log b \left( \frac{n}{\rho k} \right) + x \right) \right\} \xrightarrow{p} 0, \quad (4.26)
\]

and

\[
\sqrt{k} (\log X_{(k+1)} - \log b(n/k)) = O_P(1). \quad (4.27)
\]

Let \( \varphi(x) = \phi(x) + \frac{1}{\gamma^{1+\epsilon}} : \varphi(0) = 0. \) Then,

\[
\frac{1}{\sqrt{k}} \sum_{i=1}^{n} \varphi(Y_{(i)}) - \frac{1}{\sqrt{k}} \sum_{i=1}^{n} \varphi(Y_i) I(Y_i > 0) = o_P(1). \quad (4.28)
\]

**Proof.** Like in the proofs of Theorem 2.2 of [13] and Proposition 2.1 of [18], it can be seen that (4.26) implies

\[
\log X_{(\rho k)} - \log b \left( \frac{n}{\rho k} \right) \xrightarrow{p} 0 \quad \text{for each } \rho \text{ in } I. \quad (4.29)
\]

For \( \epsilon > 0 \) satisfying \((1 - \epsilon, 1 + \epsilon) \subset I\), we rewrite the left hand side of (4.28) as follows:

\[
\frac{1}{\sqrt{k}} \sum_{i \in I_1} \varphi(Y_{(i)}) I(Y_{(i)} \leq 0) + \frac{1}{\sqrt{k}} \sum_{i \in I_2} \varphi(Y_{(i)}) I(Y_{(i)} \leq 0) - \frac{1}{\sqrt{k}} \sum_{i \in I_3} \varphi(Y_{(i)}) I(Y_{(i)} > 0), \quad (4.30)
\]

where \( I_1 := \{1, \ldots, [(1-\epsilon)k]\}, I_2 := \{[(1-\epsilon)k] + 1, \ldots, k\}, I_3 := \{k + 1, \ldots, [(1+\epsilon)k]\}, \) and \( I_4 := \{[(1+\epsilon)k] + 1, \ldots, n\}. \) By using Taylor’s theorem, we can rewrite the second term in (4.30) as follows:

\[
\frac{1}{\sqrt{k}} \sum_{i \in I_2} \left\{ \varphi'(0) Y_{(i)} + \frac{\varphi''(\xi_{ni})}{2} Y_{(i)}^2 \right\} I(Y_{(i)} \leq 0)
\]

\[
= \frac{\varphi'(0)}{\sqrt{k}} \sum_{i \in I_2} Y_{(i)} I(Y_{(i)} \leq 0) + \frac{1}{\sqrt{k}} \sum_{i \in I_2} \frac{\varphi''(\xi_{ni})}{2} Y_{(i)}^2 I(Y_{(i)} \leq 0), \quad (4.31)
\]

where \( \xi_{ni} \) lies between \( Y_{(i)} \) and 0 for each \( i \in I_2. \) The second term in (4.31) is dominated by

\[
\sqrt{k} |\log X_{(k+1)} - \log b(n/k)| \frac{1}{k} \sum_{i \in I_2} \frac{\varphi''(\xi_{ni})}{2},
\]

which is negligible since \( \min_{i \in I_2} \xi_{ni} \) is greater than \(-1\) with probability tending to 1, \( \varphi'' \) is bounded on \([-1, \infty)\), and (4.27) holds. A similar argument can be applied to the third term in (4.30). This enables us to rewrite (4.30) as follows:
\[
\frac{1}{\sqrt{k}} \left( \sum_{i \in I_1} \varphi(Y_{i(1)}) I(Y_{i(1)} \leq 0) - \sum_{i \in I_4} \varphi(Y_{i(1)}) I(Y_{i(1)} > 0) \right) \\
+ \frac{\varphi'(0)}{\sqrt{k}} \left( \sum_{i \in I_2} Y_{i(1)} I(Y_{i(1)} \leq 0) - \sum_{i \in I_3} Y_{i(1)} I(Y_{i(1)} > 0) \right) + o_P(1). \quad (4.32)
\]

The remaining part of the proof is essentially the same as that of Lemma 2.1 of [13], and we only give a guideline. Due to (4.29), it can be shown that the first term in (4.32) is \(o_P(1)\) for each \(\epsilon\). For handling the second term, note that it is dominated by \(2\varphi'(0)\epsilon \sqrt{k}(\log X_{(k+1)} - \log b(n/k))\). By letting \(\epsilon \to 0\), (4.28) is established. \(\square\)

**Theorem 3.** Suppose that (1.3) and A hold, and there exists a random vector \((Z_1, Z_2)^T\) such that

\[
\begin{pmatrix}
\frac{1}{\sqrt{k}} \sum_{i=1}^{n} \{ \phi(Y_i) I(Y_i > 0) - E\phi(Y_i) I(Y_i > 0) \} \\
\frac{\gamma}{\sqrt{k}} \sum_{i=1}^{n} \{ Y_i^{(0)} - EY_i^{(0)} \} \\
\frac{\gamma}{\sqrt{k}} \sum_{i=1}^{n} \{ Y_i^{(\xi)} - EY_i^{(\xi)} \}
\end{pmatrix}
\Rightarrow
denote=
\begin{pmatrix}
Z_1 \\
Z_2 \\
Z_2
\end{pmatrix}
\quad \text{for all } \zeta \in \mathbb{R}. \quad (4.33)
\]

Then,

\[
\begin{pmatrix}
\frac{1}{\sqrt{k}} \sum_{i=1}^{n} \{ \phi(Y_i) I(Y_i > 0) - E\phi(Y_i) I(Y_i > 0) \} \\
\frac{\gamma}{\sqrt{k}} \sum_{i=1}^{n} \{ Y_i^{(0)} - EY_i^{(0)} \} \\
\frac{1}{\sqrt{k}} \{ \log X_{(k+1)} - \log b(n/k) \}
\end{pmatrix}
\Rightarrow
denote=
\begin{pmatrix}
Z_1 \\
Z_2 \\
Z_2
\end{pmatrix} . \quad (4.34)
\]

Further, if (4.26) holds for each \(x \in \mathbb{R}\) and \(\rho\) in a neighborhood \(I\) of 1, then

\[-\sqrt{k}U_n(\gamma) \Rightarrow Z_1 + \frac{\alpha}{\gamma^{2+\alpha}(1+\alpha)^2} Z_2.\]

**Proof.** By using the arguments in the proof of Theorem 2.4 of [13], we can see that (4.33) implies (4.34) under A, so that

\[
\sqrt{k}(\log X_{(k+1)} - \log b(n/k)) = O_P(1). \quad (4.35)
\]

On setting \(\psi(x) := \psi(x, \gamma; \alpha)\), we can write

\[
\sqrt{k}U_n(\gamma) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \psi(\tilde{Y}(i)) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left\{ \phi(Y_{i(1)}) - \phi(\tilde{Y}(i)) \right\} + \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \psi(Y_{i(1)}).
\]

Using Taylor’s theorem, we can rewrite the first term in the right hand side of the above equation as follows:
\[
\frac{1}{\sqrt{k}} \sum_{i=1}^{k} \phi' \left( \tilde{Y}_{(i)} \right) \left( \log X_{(k+1)} - \log b(n/k) \right) + \frac{1}{2\sqrt{k}} \sum_{i=1}^{k} \phi'' \left( \xi_{ni} \right) \left( \log X_{(k+1)} - \log b(n/k) \right)^2,
\]

where the \( \xi_{ni} \)'s are between \( Y_{(i)} \) and \( \tilde{Y}_{(i)} \). Since \( \phi''(x) \) is bounded on \( x \in [-1, \infty) \) and \( \min_{1 \leq i \leq k} \xi_{ni} \) is greater than \(-1\) with probability tending to 1, it follows from (4.35) that the second term in the above argument is \( o_P(1) \). Therefore, since (4.26) implies (4.19), we have

\[
\sqrt{k} U_n(\gamma) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \psi \left( Y_{(i)} \right) + \sqrt{k} c_\alpha \left( \log X_{(k+1)} - \log b(n/k) \right) + o_P(1),
\]

where we have used the boundedness and continuity of \( \phi' \) on \([0, \infty)\) and the fact that

\[
c_\alpha = \int_0^\infty \phi'(x) x^{-\frac{3}{2}} e^{-\frac{x}{2}} dx = \frac{1 + \alpha + \alpha^2}{\sqrt{2\pi} (1 + \alpha)^{\frac{3}{2}}}.
\]

Let \( \varphi(x) := \phi(x) + \frac{1}{\gamma^{1+\alpha}} \); note that \( \varphi(0) = 0 \). Then, we have from (4.35) and Lemma 3 that

\[
\frac{1}{\sqrt{k}} \sum_{i=1}^{k} \varphi \left( Y_{(i)} \right) - \frac{1}{\sqrt{k}} \sum_{i=1}^{n} \varphi \left( Y_i \right) I (Y_i > 0) = o_P(1).
\]

Combining this and Lemma 2, we have that under A,

\[
\frac{1}{\sqrt{k}} \sum_{i=1}^{k} \psi \left( Y_{(i)} \right) = -\frac{1}{\sqrt{k}} \sum_{i=1}^{k} \varphi \left( Y_{(i)} \right) + \sqrt{k} v_1(\gamma; \alpha) + \frac{\sqrt{k}}{\gamma^{1+\alpha}}
\]

\[
= -\frac{1}{\sqrt{k}} \sum_{i=1}^{n} \left\{ \varphi \left( Y_i \right) I (Y_i > 0) - \mathbb{E} \varphi \left( Y_i \right) I (Y_i > 0) \right\}
\]

\[
= -\frac{1}{\gamma^{2+\alpha}} \left( \frac{\gamma}{\sqrt{k}} \sum_{i=1}^{n} \left\{ I (Y_i > 0) - \mathbb{E} I (Y_i > 0) \right\} \right) + o_P(1).
\]

Therefore, it follows from (4.36) that under A,

\[
\sqrt{k} U_n(\gamma) = -\frac{1}{\sqrt{k}} \sum_{i=1}^{n} \left\{ \varphi \left( Y_i \right) I (Y_i > 0) - \mathbb{E} \varphi \left( Y_i \right) I (Y_i > 0) \right\}
\]

\[
-\frac{1}{\gamma^{2+\alpha}} \left( \frac{\gamma}{\sqrt{k}} \sum_{i=1}^{n} \left\{ I (Y_i > 0) - \mathbb{E} I (Y_i > 0) \right\} \right)
\]

\[+
\sqrt{k} c_\alpha \left( \log X_{(k+1)} - \log b(n/k) \right) + o_P(1).
\]

This asserts the theorem. \( \Box \)

In what follows, we verify that under some regularity conditions, \( U_n(\gamma) \) is asymptotically normal.

**Theorem 4.** Suppose that (1.3) holds, and there exists a sequence of positive integers \( \{r_n\} \) satisfying \( r_n^2 = o(n) \) (\( r_n \to \infty \)) and (2.10). Suppose that for each \( 0 < \varepsilon < 1 \) and \( \zeta \in \mathbb{R} \), there exist constants \( \chi, \vartheta, \) and \( \omega \) satisfying (2.12)–(2.15). Further, suppose that for any \( \tau > 0 \),
Proof. Let $c_1$, $c_2$, $c_3$ and $\zeta$ be any fixed real numbers. For $i = 1, \ldots, m_n$ ($m_n = [n/r_n]$) and $0 < \epsilon < 1$, we define

\[
\mathcal{J}_{ni} := \mathcal{J}_{ni}(\epsilon) = \{(i - 1)r_n + 1, \ldots, (i - 1)r_n + [(1 - \epsilon)r_n]\},
\]
\[
\mathcal{J}'_{ni} := \mathcal{J}'_{ni}(\epsilon) = \{(i - 1)r_n + [(1 - \epsilon)r_n] + 1, \ldots, ir_n\},
\]
\[
\mathcal{J}_n := \{m_nr_n + 1, \ldots, n\},
\]
and

\[
B_{ni} := \frac{1}{\sqrt{k}} \sum_{j \in \mathcal{J}_{ni}} \left\{ c_1 \left( \phi(Y_j)Y_j^{(0)} - \mathbb{E}\phi(Y_j)Y_j^{(0)} \right) + c_2 \gamma \left( Y_j^{(0)} - \mathbb{E}Y_j^{(0)} \right) + c_3 \gamma \left( Y_j^{(\zeta)} - \mathbb{E}Y_j^{(\zeta)} \right) \right\},
\]
\[
B'_{ni} := \frac{1}{\sqrt{k}} \sum_{j \in \mathcal{J}'_{ni}} \left\{ c_1 \left( \phi(Y_j)Y_j^{(0)} - \mathbb{E}\phi(Y_j)Y_j^{(0)} \right) + c_2 \gamma \left( Y_j^{(0)} - \mathbb{E}Y_j^{(0)} \right) + c_3 \gamma \left( Y_j^{(\zeta)} - \mathbb{E}Y_j^{(\zeta)} \right) \right\},
\]
\[
R_n := \frac{1}{\sqrt{k}} \sum_{j \in \mathcal{J}_n} \left\{ c_1 \left( \phi(Y_j)Y_j^{(0)} - \mathbb{E}\phi(Y_j)Y_j^{(0)} \right) + c_2 \gamma \left( Y_j^{(0)} - \mathbb{E}Y_j^{(0)} \right) + c_3 \gamma \left( Y_j^{(\zeta)} - \mathbb{E}Y_j^{(\zeta)} \right) \right\}.
\]
From (2.10), we have
\[
\left| \mathbb{E} \exp \left\{ i \sum_{i=1}^{m_n} B_{ni} \right\} - \prod_{1 \leq i \leq m_n} \mathbb{E} \exp \{ i B_{ni} \} \right| \leq 16(m_n - 1)\beta([\epsilon r_n]) \quad \text{for all real } t,
\]
where \( i \) is the imaginary unit (cf. Lemma 2 of [5, P. 365]). Thus, owing to (2.10), we can view \( \sum_{i=1}^{m_n} B_{ni} \) as a sum of i.i.d. copies of \( B_{n1} \). Let \( c = (c_1, c_2, c_3)^T \). Using (2.12)–(2.15) and Lemma 2, we have
\[
\lim_{n \to \infty} m_n \text{Var}(B_{n1}) = (1 - \epsilon) c^T \Sigma c
\]
and
\[
\lim_{n \to \infty} m_n \text{Var}(B'_{n1}) = \epsilon c^T \Sigma c.
\]
Therefore, by (4.38),
\[
\sum_{i=1}^{m_n} B_{ni} \Rightarrow N(0, (1 - \epsilon)c^T \Sigma c) \quad \text{(4.40)}
\]
and
\[
\sum_{i=1}^{m_n} B'_{ni} \Rightarrow N(0, \epsilon c^T \Sigma c). \quad \text{(4.41)}
\]
On the other hand, due to the fact that \( r_n^2 = o(n) \) and Lemma 2,
\[
\text{Var}(R_n) \leq \frac{r_n^2}{n} \frac{2n}{k} \left( \text{Var}(\phi(Y_1)Y_1^{(0)}) + \text{Var}(Y_1^{(0)}) + \text{Var}(Y_1^{(\xi)}) \right) \to 0 \quad \text{as } n \to \infty,
\]
which implies that \( R_n \) is \( o_P(1) \). By combining this, (4.40) and (4.41), and letting \( \epsilon \to 0 \), we have
\[
\frac{1}{\sqrt{k}} \sum_{j=1}^{n} \left\{ c_1 \left( \phi(Y_j)Y_j^{(0)} - \mathbb{E}\phi(Y_j)Y_j^{(0)} \right) + c_2 Y_j^{(0)} - \mathbb{E}Y_j^{(0)} \right\} + c_3 Y_j^{(\xi)} - \mathbb{E}Y_j^{(\xi)}
\]
\[
\Rightarrow N(0, c^T \Sigma c).
\]
This completes the proof. \( \square \)

We can relax the conditions in Theorem 4 by assuming more stringent strong mixing conditions. The result is as follows.

**Corollary 2.** Suppose that (1.3) and **A** hold, and there exists a sequence of positive integers \( \{r_n\} \) satisfying \( r_n \to \infty, r_n^2 = o(k) \), and (2.10). Suppose that for each \( 0 < \epsilon < 1 \), there exist constants \( \chi, \vartheta, \) and \( \omega \) such that (2.12), (2.13) and (2.15) hold with \( \zeta = 0 \). Then, the same result as in Theorem 4 holds.

**Proof.** Since for each \( c_1, c_2, \) and \( c_3 \), the summands of \( W_n \) in Theorem 4 have a common bound not depending on \( n \), (4.38) holds for each \( c_1, c_2, c_3, \zeta, \) and \( \tau > 0 \). Thus, it suffices to show that for any real number \( \zeta \),
\[
\frac{1}{\sqrt{k}} \sum_{i=1}^{n} \left\{ Y_i^{(0)} - Y_i^{(\xi)} - \mathbb{E} \left( Y_1^{(0)} - Y_1^{(\xi)} \right) \right\} = o_P(1).
\]
Without loss of generality, we can assume that \( m_n \beta(r_n) \to 0 \). Here, we express the left hand side of the above equation as \( \sum_{i=1}^{m_n} S_{ni} + R_n \) where

\[
S_{ni} := \frac{1}{\sqrt{k}} \sum_{j=(i-1)r_n+1}^{ir_n} \left\{ Y_j^{(0)} - Y_j^{(\zeta)} - \mathbb{E} \left( Y_1^{(0)} - Y_1^{(\zeta)} \right) \right\}
\]

and

\[
R_n := \frac{1}{\sqrt{k}} \sum_{j=mnr_n+1}^{n} \left\{ Y_j^{(0)} - Y_j^{(\zeta)} - \mathbb{E} \left( Y_1^{(0)} - Y_1^{(\zeta)} \right) \right\},
\]

which is \( o_P(1) \). Now, if we set \( O_n := \{ i : i \text{ is odd in } 1, \ldots, m_n \} \), we can view \( \{ S_{ni} : i \in O_n \} \) as i.i.d. random variables since \( \lim_{n \to \infty} m_n \beta(r_n) = 0 \). Thus, for any \( \epsilon > 0 \),

\[
P \left( \left| \sum_{i \in O_n} S_{ni} \right| > \epsilon \right) \leq \frac{1}{\epsilon^2} \sum_{i \in O_n} \text{Var} S_{ni} \leq \frac{nr_n}{\epsilon^2k} \mathbb{E} \left| Y_1^{(0)} - Y_1^{(\zeta)} \right| = o(1),
\]

where we have used the fact that \( n \mathbb{E} \left| Y_1^{(0)} - Y_1^{(\zeta)} \right| = O(\sqrt{k}) \) which is due to Lemma 2. Since a similar argument can be applied to \( E_n := \{ 1, \ldots, m_n \} \setminus O_n \), the corollary is established. \( \square \)

**Proof of Theorem 2.** From (2.11), there exists a sequence \( \{ \hat{\gamma}_n \} \) satisfying \( U_n(\hat{\gamma}_n) = 0 \) such that \( \hat{\gamma}_n \overset{P}{\to} \gamma \). By using Taylor’s theorem, we can write

\[
-\sqrt{k} U_n(\gamma) = \sqrt{k}(\hat{\gamma}_n - \gamma) \left\{ \hat{U}_n(\gamma) + o_P(1) \right\}.
\]

Since \( \hat{U}_n(\gamma) \overset{P}{\to} J(\gamma; \alpha) \), the asymptotic normality of \( -\sqrt{k} U_n(\gamma) \) implies that of \( \sqrt{k}(\hat{\gamma}_n - \gamma) \). The asymptotic normality of \( -\sqrt{k} U_n(\gamma) \) follows from Theorems 3 and 4. This completes the proof. \( \square \)

**Acknowledgments**

We thank an AE and the two referees for their valuable comments leading to improvement of the quality of this paper. This work was supported by grant No. R01-2006-000-10545-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

**References**