Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint

On maximum likelihood estimation of the extreme value index

Holger Drees, Ana Ferreira and Laurens de Haan

Preprint No. 72

Saarbrücken 2002

Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

On maximum likelihood estimation of the extreme value index

Holger Drees

Universität des Saarlandes
Fachbereich Mathematik
Postfach 15 11 50
66041 Saarbrücken
Germany
E-Mail: drees@num.uni-sb.de

Ana Ferreira

Departamento de Matematica
Instituto Superior de Agronomia
Tapada da Ajuda
1349-017 Lisboa
Portugal
E-Mail: anafh@madpet.isa.utl.pt

Laurens de Haan

Econometrisch Instituut
Faculteit der Economische Wetenschappen
Erasmus Universiteit Rotterdam
PO Box 1738
3000 DR Rotterdam
The Netherlands
E-Mail: ldhaan@few.eur.nl

submitted: October 24, 2002

Preprint No. 72

Saarbrücken 2002

Edited by
FR 6.1 – Mathematik
Im Stadtwald
D–66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Abstract

We prove asymptotic normality of the so-called maximum likelihood estimator of the extreme value index.

AMS Subject Classification: 62G32, 62G20
Key words: asymptotic normality, exceedances, extreme value index, maximum likelihood, order statistics, second order condition

1 Introduction

Let $X_1, X_2, \ldots$ be independent and identically distributed (i.i.d.) random variables (r.v.'s) from some unknown distribution function (d.f.) F. Denote the upper endpoint of F by x^*, where $x^* = \sup\{x : F(x) < 1\} \leq \infty$, and let

$$F_t(x) = P(X \leq t + x | X > t) = \frac{F(t+x) - F(t)}{1 - F(t)}, \tag{1}$$

with $1 - F(t) > 0$, $t < x^*$ and $x > 0$, be the conditional d.f. of $X - t$ given $X > t$. Then it is well known (Balkema and de Haan, 1974; Pickands, 1975) that up to scale and location transformations the generalized Pareto d.f. given by

$$H_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}, \tag{2}$$

$x > 0$ if $\gamma \geq 0$ and $0 < x < -1/\gamma$ if $\gamma < 0$ (for $\gamma = 0$ read $(1 + \gamma x)^{-1/\gamma}$ as $\exp(-x)$), can provide a good approximation to the conditional probabilities (1). More precisely, it has been proved that there exists a normalizing function $\sigma(t) > 0$, such that

$$\lim_{t \to x^*} F_t(x\sigma(t)) \to H_\gamma(x)$$

for all x, or equivalently

$$\lim_{t \to x^*} \sup_{0 < x < x^* - t} |F_t(x) - H_\gamma(x/\sigma(t))| = 0 \tag{3}$$

if and only if F is in the maximum domain of attraction of the corresponding extreme value d.f. $G_\gamma(x) = \exp(-(1 + \gamma x)^{-1/\gamma})$ (Gnedenko, 1943), commonly denoted by $F \in D(G_\gamma)$.

Under this set-up, it turns out that a major issue for estimating extreme events is the estimation of the extreme value index γ. A variety of procedures to estimate γ are now available in the literature (e.g. Hill, 1975; Dekkers et al., 1989; Smith, 1987), although there are still open problems. Quite often the accuracy of these estimators rely heavily on the choice of some threshold but, it is not our aim here to address this type of optimality questions.

The so-called maximum likelihood estimator (m.l.e.) is one of the most popular estimators in this context. For a sample of size n let $X_{1,n} \leq X_{2,n} \leq \ldots X_{n,n}$ be the ascending order statistics. In view of the generalized Pareto approximation, it is clear that we shall base our inference on some high order statistics, say $(X_{n-k,n}, X_{n-k+1,n}, \ldots, X_{n,n})$. Consider the one-to-one mapping

$$\begin{cases} Y_0 & = X_{n-k,n} \\ Y_1 & = X_{n-k+1,n} - X_{n-k,n} \\ \cdots & \\ Y_k & = X_{n,n} - X_{n-k,n}. \end{cases}$$

Any inference based on maximizing the likelihood of $(X_{n-k,n}, X_{n-k+1,n}, \ldots, X_{n,n})$ is equivalent to considering the likelihood of $(Y_0, Y_1, \ldots, Y_k)$. Now the distribution of $(Y_1, \ldots, Y_k)$ given $Y_0 = y_0$ equals the distribution of $(Y^*_{1,k}, \ldots, Y^*_{k,k})$, the order statistics of an i.i.d. sample $(Y^*_1, \ldots, Y^*_k)$ with common distribution $F_{y_0}(x) = P(X \leq y_0 + x | X > y_0)$ (see e.g. Theorem 2.4.1 in Arnold et al., 1992). Therefore, the common approach is first to factorize the distribution of $(Y^*_1, \ldots, Y^*_k)$ and to use the generalized Pareto approximation (3), and second to ignore the factor related to the marginal distribution of Y_0 in the likelihood of $(Y_0, Y_1, \ldots, Y_k)$. To sum up, the so-called m.l.e.'s of γ and σ are obtained by maximizing in γ and σ the approximative likelihood $\prod_{i=1}^k h_{\gamma,\sigma}(Y_1, \ldots, Y_k)$ where $h_{\gamma,\sigma}(y) = \partial H_\gamma(y/\sigma)/\partial y$. Note that the approximative likelihood function tends to ∞ if $\gamma < -1$ and $\gamma/\sigma \downarrow -1/(X_{n,n} - X_{n-k,n})$, so that a maximum over the full range of possible values for (γ, σ) does not exist. Since moreover the maximum likelihood estimator behaves irregular if $\gamma \leq -1/2$, we look for a maximum of the approximative likelihood function only in the region $(\gamma, \sigma) \in (-1/2, \infty) \times (0, \infty)$.

The likelihood equations are then given in terms of the partial derivatives

$$\begin{cases} \dfrac{\partial \log h_{\gamma,\sigma}(y)}{\partial \gamma} = \dfrac{1}{\gamma^2} \log\left(1 + \dfrac{\gamma}{\sigma} y\right) - \left(\dfrac{1}{\gamma} + 1\right) \dfrac{\frac{y}{\sigma}}{1 + \frac{\gamma}{\sigma} y}, \\ \dfrac{\partial \log h_{\gamma,\sigma}(y)}{\partial \sigma} = -\dfrac{1}{\sigma} - \left(\dfrac{1}{\gamma} + 1\right) \dfrac{-\frac{\gamma}{\sigma^2} y}{1 + \frac{\gamma}{\sigma} y}, \end{cases}$$

where for $\gamma = 0$ these terms should be interpreted as

$$\begin{cases} \dfrac{\partial \log h_{\gamma,\sigma}(y)}{\partial \gamma}\Big|_{\gamma=0} = \dfrac{1}{2}\left(\dfrac{y}{\sigma}\right)^2 - \dfrac{y}{\sigma}, \\ \dfrac{\partial \log h_{\gamma,\sigma}(y)}{\partial \sigma}\Big|_{\gamma=0} = -\dfrac{1}{\sigma} + \dfrac{y}{\sigma^2} = 0. \end{cases}$$

The resulting likelihood equations in terms of the excesses $X_{n-i+1,n} - X_{n-k,n}$ are as follows

$$\begin{cases} \displaystyle\sum_{i=1}^k \frac{1}{\gamma^2} \log\left(1 + \frac{\gamma}{\sigma}(X_{n-i+1,n} - X_{n-k,n})\right) \\ \qquad\qquad - \left(\dfrac{1}{\gamma} + 1\right) \dfrac{\frac{1}{\sigma}(X_{n-i+1,n} - X_{n-k,n})}{1 + \frac{\gamma}{\sigma}(X_{n-i+1,n} - X_{n-k,n})} = 0 \\ \displaystyle\sum_{i=1}^k \left(\frac{1}{\gamma} + 1\right) \frac{\frac{\gamma}{\sigma}(X_{n-i+1,n} - X_{n-k,n})}{1 + \frac{\gamma}{\sigma}(X_{n-i+1,n} - X_{n-k,n})} = k \end{cases} \tag{4}$$

(with a similar interpretation when $\gamma = 0$), which for $\gamma \neq 0$ can be simplified to

$$\begin{cases} \frac{1}{k}\sum_{i=1}^{k} \log\left(1+\frac{\gamma}{\sigma}(X_{n-i+1,n}-X_{n-k,n})\right) = \gamma \\ \frac{1}{k}\sum_{i=1}^{k} \frac{1}{1+\frac{\gamma}{\sigma}(X_{n-i+1,n}-X_{n-k,n})} = \frac{1}{\gamma+1}, \end{cases}$$

with $(\gamma, \sigma) \in (-1/2, \infty) \times (0, \infty)$. The numerical problem to find a solution of these equations which maximizes the approximative likelihood was discussed by Grimshaw (1993).
From the above reasoning it follows that the m.l.e. of γ is shift and scale invariant, and the m.l.e. of σ is shift invariant and scale equivariant.
Proofs of the asymptotic normality of the m.l.e.'s of γ and σ were given by Smith (1987), and also by Drees (1998) in the case $\gamma > 0$. Nonetheless we consider some proofs not easily understandable. Moreover, some of the conditions used in the aforementioned papers are unnecessarily restrictive. In this paper we present a relatively simple direct approach to prove asymptotic normality of the m.l.e.'s of γ and σ. It is based on some recent approximations to the tail empirical quantile function established by Drees (1998).
Next we sketch the proof of the asymptotic normality. Under standard second order conditions we have for an intermediate sequence $k_n \in \mathbb{N}$, $n \in \mathbb{N}$, (i.e. $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$)

$$\left(\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)_{t\in[0,1]} = \left(\frac{t^{-\gamma_0}-1}{\gamma_0} + k_n^{-1/2}Y_n(t)\right)_{t\in[0,1]} \tag{5}$$

where $(Q_n(t))_{t\in[0,1]}$ is a distributionally equivalent version of the process $(X_{n-[k_n t],n})_{t\in[0,1]}$, $(Y_n(t))_{t\in[0,1]}$ is an asymptotically Gaussian process of known mean and covariance function (Lemma 3.1), γ_0 is the true parameter and $\tilde{a}$ is a suitably chosen positive function. Hence for all $t \in [0,1]$

$$t^{\gamma_0}\left(1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right) = 1 + \left(\frac{\gamma}{\tilde{\sigma}}-\gamma_0\right)\frac{1-t^{\gamma_0}}{\gamma_0} + t^{\gamma_0}\frac{\gamma}{\tilde{\sigma}}k_n^{-1/2}Y_n(t) \tag{6}$$

where $\tilde{\sigma} = \sigma/\tilde{a}(\frac{k_n}{n})$. Now if the sequence of solutions $(\gamma, \tilde{\sigma})$ satisfies

$$\gamma - \gamma_0 = O_p(k_n^{-1/2}) \quad \text{and} \quad \tilde{\sigma} - 1 = O_p(k_n^{-1/2}),$$

one can prove, using a construction similar to (5), that

$$\inf_{1/(2k_n)\le t\le 1} t^{\gamma_0}\left(1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)$$

is stochastically bounded away from 0 (Lemma 3.2). This implies

$$\begin{aligned}
&\log\left(t^{\gamma_0}\left(1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)\right)\\
&= \log\left(1+\left(\frac{\gamma}{\tilde\sigma}-\gamma_0\right)\frac{1-t^{\gamma_0}}{\gamma_0}+t^{\gamma_0}\frac{\gamma}{\tilde\sigma}k_n^{-1/2}Y_n(t)\right)\\
&= \left(\frac{\gamma}{\tilde\sigma}-\gamma_0\right)\frac{1-t^{\gamma_0}}{\gamma_0}+t^{\gamma_0}\frac{\gamma}{\tilde\sigma}k_n^{-1/2}Y_n(t)+o_p(k_n^{-1/2})
\end{aligned}$$

and

$$\frac{1}{1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}}=t^{\gamma_0}\left(1-\left(\frac{\gamma}{\tilde\sigma}-\gamma_0\right)\frac{1-t^{\gamma_0}}{\gamma_0}-t^{\gamma_0}\frac{\gamma}{\tilde\sigma}k_n^{-1/2}Y_n(t)+o_p(k_n^{-1/2})\right)$$

where the o_p-term is uniform for $1/(2k_n)\le t\le 1$ (proof of Proposition 3.1). Hence, up to a $o_p(k_n^{-1/2})$-term the equations (4) are equivalent to linear equations which can be solved readily. The proof in case $\gamma_0=0$ requires longer expansions but is similar.
The precise statement about the asymptotic normality is given in Theorem 2.1. In Theorem 2.2 an equivalent explicit estimator is constructed in the case $\gamma_0=0$.
Throughout the paper, $F^{\leftarrow}$ denotes the generalized inverse of F, $\to_d$ convergence in distribution and $\to_p$ convergence in probability.

2 Asymptotic normality of the maximum likelihood estimators

Assume that there exist measurable, locally bounded functions $a,\Phi:(0,1)\to(0,\infty)$ and $\Psi:(0,\infty)\to\mathbb{R}$ such that

$$\lim_{t\downarrow 0}\frac{\frac{F^{\leftarrow}(1-tx)-F^{\leftarrow}(1-t)}{a(t)}-\frac{x^{-\gamma_0}-1}{\gamma_0}}{\Phi(t)}=\Psi(x), \tag{7}$$

for some $\gamma_0>-1/2$, for all $t\in(0,1)$ and $x>0$, where $x\mapsto\Psi(x)/(x^{-\gamma_0}-1)$ is not constant, $\Phi(t)$ not changing sign eventually and $\Phi(t)\to 0$ as $t\downarrow 0$. Then, according to de Haan and Stadtmüller (1996), $|\Phi|$ is $-\rho$-varying at 0 for some $\rho\le 0$, i.e. $\lim_{t\downarrow 0}\Phi(tx)/\Phi(t)=x^{-\rho}$ for all $x>0$, and

$$\Psi(x)=\begin{cases}(x^{-(\gamma_0+\rho)}-1)/(\gamma_0+\rho) & ,\rho<0\\ -x^{-\gamma_0}\log(x)/\gamma_0 & ,\gamma_0\neq\rho=0\\ \log^2(x) & ,\gamma_0=\rho=0,\end{cases} \tag{8}$$

provided that the normalizing function a and the function Φ are chosen suitably. Condition (7) is a second order refinement of $F\in D(G_{\gamma_0})$. Still, it is a quite general condition, satisfied for all usual distributions satisfying the max-domain of attraction condition.
We assume throughout that k_n is an intermediate sequence, i.e. $k_n\to\infty$ and $k_n/n\to 0$ as $n\to\infty$.

Theorem 2.1. *Assume condition* (7) *for some* $\gamma_0 > -1/2$ *and that the intermediate sequence* k_n *satisfies*

$$\Phi(k_n/n) = O(k_n^{-1/2}). \tag{9}$$

Then the system of likelihood equations (4) *has a sequence of solutions* $(\hat\gamma_n, \hat\sigma_n)$ *that verifies*

$$\begin{aligned} &k_n^{1/2}(\hat\gamma_n - \gamma_0) - \frac{(\gamma_0+1)^2}{\gamma_0} k_n^{1/2}\Phi(\frac{k_n}{n}) \int_0^1 \left(t^{\gamma_0} - (2\gamma_0+1)t^{2\gamma_0}\right)\Psi(t)dt \\ &\quad \to_d \ \frac{(\gamma_0+1)^2}{\gamma_0} \int_0^1 \left(t^{\gamma_0} - (2\gamma_0+1)t^{2\gamma_0}\right)\left(W(1) - t^{-(\gamma_0+1)}W(t)\right)\, dt \end{aligned} \tag{10}$$

$$\begin{aligned} &k_n^{1/2}\left(\frac{\hat\sigma_n}{a(\frac{k_n}{n})} - 1\right) - \frac{\gamma_0+1}{\gamma_0} k_n^{1/2}\Phi(\frac{k_n}{n}) \int_0^1 \left((\gamma_0+1)(2\gamma_0+1)t^{2\gamma_0} - t^{\gamma_0}\right)\Psi(t)\, dt \\ &\quad \to_d \ \frac{\gamma_0+1}{\gamma_0} \int_0^1 \left((\gamma_0+1)(2\gamma_0+1)t^{2\gamma_0} - t^{\gamma_0}\right)\left(W(1) - t^{-(\gamma_0+1)}W(t)\right) dt \end{aligned} \tag{11}$$

as $n \to \infty$, *and the convergence holds jointly with the same standard Brownian motion* W. *For* $\gamma_0 = 0$ *these equations should be interpreted as their limits when* $\gamma_0 \to 0$, *i.e.*

$$\begin{aligned} &k_n^{1/2}\hat\gamma_n + k_n^{1/2}\Phi(\frac{k_n}{n}) \int_0^1 (2+\log t)\Psi(t)dt \\ &\quad \to_d \ -\int_0^1 (2+\log t)\left(W(1) - t^{-1}W(t)\right)\, dt \end{aligned} \tag{12}$$

$$\begin{aligned} &k_n^{1/2}\left(\frac{\hat\sigma_n}{a(\frac{k_n}{n})} - 1\right) - k_n^{1/2}\Phi(\frac{k_n}{n}) \int_0^1 (3+\log t)\Psi(t)dt \\ &\quad \to_d \ \int_0^1 (3+\log t)\left(W(1) - t^{-1}W(t)\right)\, dt. \end{aligned} \tag{13}$$

Moreover, any sequence of solutions $(\hat\gamma_n^*, \hat\sigma_n^*)$ *which is not of the type* (10)–(13) *must satisfy* $k_n^{1/2}|\hat\gamma_n^* - \gamma_0| \to_p \infty$ *or* $k_n^{1/2}|\hat\sigma_n^*/a(k_n/n) - 1| \to_p \infty$.

Remark 2.1. Condition (9) is satisfied if $k_n \to \infty$ not too fast. The bias term $(\gamma_0+1)^2 k_n^{1/2}\Phi(k_n/n)\int_0^1 \left(t^{\gamma_0} - (2\gamma_0+1)t^{2\gamma_0}\right)\Psi(t)dt/\gamma_0$ in (10) vanishes if $k_n^{1/2}\Phi(k_n/n) \to 0$. A similar remark applies to (11)-(13).

Remark 2.2. Note that the likelihood equations are satisfied with $\gamma = 0$ if and only if

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

and $\sigma = \sum_{i=1}^{k}(X_{n-i+1,n} - X_{n-k,n})/k$. Hence the m.l.e. for γ will a.s. not be equal to 0 if, e.g., F possesses a density.

Corollary 2.1. *Under the conditions of Theorem 2.1 and if*

$$k_n^{1/2}\Phi(\frac{k_n}{n}) \to \lambda \in \mathbb{R}, \tag{14}$$

the solutions (10)-(13) *verify*

$$k_n^{1/2}\left[\begin{array}{c} \hat{\gamma}_n - \gamma_0 \\ \hat{\sigma}_n/a(k_n/n) - 1 \end{array}\right] \to_d N(\lambda\mu, \Sigma), \tag{15}$$

where N denotes the bivariate normal distribution, μ equals

$$\begin{array}{cl} \left[\frac{\rho(\gamma_0+1)}{(1-\rho)(\gamma_0-\rho+1)}, \frac{1-2\rho+\gamma_0-\rho\gamma_0}{(1-\rho)(\gamma_0-\rho+1)}\right]^T & , \textit{if } \rho < 0, \\ \left[1, \gamma_0^{-1}\right]^T & , \textit{if } \gamma_0 \neq \rho = 0, \\ \left[2, 0\right]^T & , \textit{if } \gamma_0 = \rho = 0 \end{array}$$

and

$$\Sigma = \left[\begin{array}{cc} (1+\gamma_0)^2 & -(1+\gamma_0) \\ -(1+\gamma_0) & 2+2\gamma_0+\gamma_0^2 \end{array}\right].$$

Remark 2.3. Smith (1987) examined a slightly different version of the m.l.e. that is based on the excesses over a *deterministic* threshold $u = u_n$ instead of the excesses over the random threshold $X_{n-k_n,n}$. For the comparison of Smith's results with Corollary 2.1, we focus on the case $\gamma_0 \neq 0$, $\rho < 0$ and $\lambda = 0$, when there is no asymptotic bias, since in the other cases the more restrictive second order conditions used by Smith are not directly comparable to our setting.
Let K denote the (random) number of exceedances over the threshold u and let

$$\sigma_n = \left\{\begin{array}{ll} \gamma_0 u, & \gamma_0 > 0, \\ |\gamma_0|(F^{\leftarrow}(1) - u), & \gamma_0 < 0. \end{array}\right.$$

Then it was shown that the standardized m.l.e.'s $K^{1/2}(\hat{\gamma}_n - \gamma_0, \hat{\sigma}_n/\sigma_n - 1)$ based on the exceedances $X_i - u$ converge to a centered bivariate normal distribution with covariance matrix

$$\left(\begin{array}{cc} (1+\gamma_0)^2 & -(1+\gamma_0) \\ -(1+\gamma_0) & 2(1+\gamma_0) \end{array}\right).$$

At first glance, it seems peculiar that we obtain a different asymptotic variance for the scale estimator in Corollary 2.1, namely $2(1+\gamma_0)+\gamma_0^2$. However, the following heuristic reasoning shows that in fact the increase in the variance is due to the slightly different standardization.
To make the results about the asymptotic behavior comparable, in our setting one has to condition at the event $X_{n-k_n,n} = u$. Then Smith's result claims that conditionally $\hat{\sigma}_n = \sigma_n(1 + k_n^{-1/2} Z_n)$ for some asymptotically centered normal r.v. Z_n with asymptotic variance $2(1+\gamma_0)$. Hence conditionally at $X_{n-k_n,n} = u$

$$k^{1/2}\left(\frac{\hat{\sigma}_n}{a(k_n/n)} - 1\right) = \frac{\sigma_n}{a(k_n/n)} Z_n + k^{1/2}\left(\frac{\sigma_n}{a(k_n/n)} - 1\right).$$

Because, in the restrictive setting considered here, $a(k_n/n) = \gamma_0 F^{\leftarrow}(1-k_n/n)$ for $\gamma_0 > 0$ and $a(k_n/n) = |\gamma_0|(F^{\leftarrow}(1) - F^{\leftarrow}(1-k_n/n))$ for $\gamma_0 < 0$, unconditionally (i.e. when u is replaced with $X_{n-k_n,n}$) $\sigma_n/a(k_n/n) \to 1$ in probability, so that the first term tends to a normal random variable with variance $2(\gamma_0+1)$. According to the approximation of the tail empirical quantile function (cf. (17)), unconditionally the second term converges to $\gamma_0 W(1)$. Since asymptotically $X_{n-k_n,n}$ and the excesses $X_{n-i+1,n} - X_{n-k_n,n}, 1 \leq i \leq k_n$ are independent, so are Z_n and $W(1)$. Hence the two variances $2(\gamma_0+1)$ and γ_0^2 add up, leading to the variance given in Corollary 2.1.

We now show that if $\gamma_0 = 0$ the m.l.e.'s are asymptotically equivalent in some sense to explicit estimators. Define

$$m_n^{(j)} = \frac{1}{k_n}\sum_{i=1}^{k_n}(X_{n-i+1,n} - X_{n-k_n,n})^j, \quad j = 1,2,$$

$$\hat{\gamma}_* = 1 - \frac{1}{2}\Big(1 - \frac{(m_n^{(1)})^2}{m_n^{(2)}}\Big)^{-1}$$

and

$$\hat{a}_*(\frac{k_n}{n}) = \frac{2(m_n^{(1)})^3}{m_n^{(2)}}.$$

It can be shown, using Corollary 3.1, that these estimators are consistent and asymptotically normal if $\gamma < 1/2$. Let $(\hat{\gamma}_{MLE}, \hat{\sigma}_{MLE})$ be a sequence of solutions of (4) as described in Theorem 2.1.

Theorem 2.2. *If F is in the class of distributions that satisfy* (7) *with $\gamma_0 = 0$ and if* (9) *holds, then*

$$k_n^{1/2}\,(\hat{\gamma}_* - \hat{\gamma}_{MLE}) \to_p 0$$

and

$$k_n^{1/2}\left(\frac{\hat{a}_*(\frac{k_n}{n}) - \hat{\sigma}_{MLE}}{a(\frac{k_n}{n})}\right) \to_p 0.$$

Remark 2.4. If, in addition, (7) holds with $\rho < 0$, $\sup\{x|F(x) < 1\} > 0$ and $k_n = o(\log^2 n)$, then we have an analogous result for the moment estimator introduced by Dekkers et al. (1989):

$$k_n^{1/2}\,(\hat{\gamma}_{MOM} - \hat{\gamma}_{MLE}) \to_p 0$$

where

$$\hat{\gamma}_{MOM} = M_n^{(1)} + 1 - \frac{1}{2}(1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}})^{-1}$$

with $M_n^{(j)} = \frac{1}{k_n}\sum_{i=0}^{k_n-1}(\log X_{n-i,n} - \log X_{n-k_n,n})^j$, $j = 1,2$. A similar statement holds for the scale estimator

$$\hat{a}_{**}(\frac{k_n}{n}) = \frac{2(M_n^{(1)})^3}{M_n^{(2)}}.$$

The condition $k_n = o(\log^2 n)$ ensures that the bias vanishes asymptotically. We prove this remark in Section 3.

3 Proofs

Given (7) with $\gamma_0 > -1/2$ and (9), from Theorem 2.1 in Drees (1998) one can find a probability space and define on that space a Brownian motion W and a sequence of stochastic processes Q_n such that: (i) for each n, $(Q_n(t))_{t\in[0,1]} =_d (X_{n-[k_n t],n})_{t\in[0,1]}$, and (ii) there exist functions $\tilde{a}(k_n/n) = a(k_n/n)(1 + o(\Phi(k_n/n)))$ and $\tilde{\Phi}(k_n/n) \sim \Phi(k_n/n)$ such that for all $\varepsilon > 0$

$$\sup_{t\in[0,1]} t^{\gamma_0+1/2+\varepsilon} \Big| \frac{Q_n(t) - F^{\leftarrow}(1-\frac{k_n}{n})}{\tilde{a}(\frac{k_n}{n})} - \left(\frac{t^{-\gamma_0}-1}{\gamma_0} - t^{-(\gamma_0+1)} \frac{W(k_n t)}{k_n} + \tilde{\Phi}(\frac{k_n}{n})\Psi(t) \right) \Big|$$
$$= o_p(k_n^{-1/2}) + o_p(\tilde{\Phi}(\frac{k_n}{n})), \quad \text{as} \quad n \to \infty.$$

A similar expansion is also valid for $\gamma_0 \le -1/2$ when $F^{\leftarrow}(1-k_n/n)$ is replaced with a suitable random variable.
Define

$$Y_n(t) = k_n^{1/2} \left(\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})} - \frac{t^{-\gamma_0}-1}{\gamma_0} \right) \tag{16}$$

(read $(t^{-\gamma_0} - 1)/\gamma_0$ as $-\log t$, when $\gamma_0 = 0$). Hence we have the following lemma.

Lemma 3.1. *Suppose* (7) *and that the intermediate sequence k_n satisfies* (9)*. Then, for all $\varepsilon > 0$*

$$Y_n(t) = W_n(1) - t^{-(\gamma_0+1)} W_n(t) + k_n^{1/2} \tilde{\Phi}(\frac{k_n}{n})\Psi(t) + o_p(1) t^{-(\gamma_0+1/2+\varepsilon)} \tag{17}$$

as $n \to \infty$, where $W_n(t) = k_n^{-1/2} W(k_n t)$ is a standard Brownian motion and the o_p-term is uniform for $t \in [0,1]$.

From this lemma the following corollary follows easily.

Corollary 3.1. *Under the conditions of Lemma 3.1, for all $\varepsilon > 0$*

$$Y_n(t) = O_p(1)\, t^{-(\gamma_0+1/2+\varepsilon)} \tag{18}$$

as $n \to \infty$, where the O_p-term is uniform for $t \in [0,1]$.

Given the previous results, to prove Theorem 2.1 it is sufficient to consider the likelihood equations with $(X_{n-[k_n t],n} - X_{n-k_n,n})$ replaced by $Q_n(t) - Q_n(1)$, $t \in [0,1]$. It is convenient to reparametrize the equations in terms of $(\gamma, \tilde{\sigma})$ $= (\gamma, \sigma/\tilde{a}(k_n/n))$. Then we have the equations

$$\begin{cases} \int_0^1 \left(\frac{1}{\gamma^2} \log\left(1 + \frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right) - \left(\frac{1}{\gamma}+1\right) \frac{\frac{1}{\tilde{\sigma}} \frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}}{1+\frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}} \right) dt == 0 \\ \int_0^1 \left(\frac{1}{\gamma}+1\right) \frac{\frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}}{1+\frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}} dt = 1. \end{cases} \tag{19}$$

Lemma 3.2. *Assume conditions* (7) *and* (9)*. Let* $(\gamma, \tilde{\sigma}) = (\gamma_n, \tilde{\sigma}_n)$ *be such that*

$$|\gamma/\tilde{\sigma} - \gamma_0| = O_p(k_n^{-1/2}). \tag{20}$$

Then, if $-1/2 < \gamma_0 < 0$ *or* $\gamma_0 > 0$,

$$P\left(1 + \frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})} \geq C_n t^{-\gamma_0}, \quad t \in [\frac{1}{2k_n}, 1]\right) \to 1, \qquad n \to \infty, \tag{21}$$

for some r.v.'s $C_n > 0$ *such that* $1/C_n = O_P(1)$*. If* $\gamma_0 = 0$,

$$P\left(1 + \frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})} \geq \frac{1}{2}, \quad t \in [\frac{1}{2k_n}, 1]\right) \to 1, \qquad n \to \infty, \tag{22}$$

and

$$\sup_{t \in [0,1]} \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})} = O_p(\log k_n), \qquad n \to \infty. \tag{23}$$

Proof. It suffices to prove the assertions with $Q_n(t)$ replaced with $X_{n-[k_n t],n}$. Without loss of generality, one may assume $X_{i,n} = F^{\leftarrow}(1 - U_{i,n})$ for uniform order statistics $U_{i,n}$ since $(X_{n-[k_n t],n})_{t \in [0,1]} =^d (F^{\leftarrow}(1 - U_{[k_n t],n})_{t \in [0,1]} =^d (Q_n(t))_{t \in [0,1]}$.
Note that by Shorack and Wellner (1986, Chapter 10, Section 3, p. 416, inequality 2)

$$\sup_{1/(2k_n) \leq t \leq 1} \frac{nU_{[k_n t]+1,n}}{k_n t} = O_P(1), \qquad \sup_{0 \leq t \leq 1} \frac{k_n t}{nU_{[k_n t]+1,n}} = O_P(1) \tag{24}$$

as $n \to \infty$. Also note that (7) implies, for some functions $\tilde{a}(s) \sim a(s)$ and $\tilde{\Phi}(s) \sim \Phi(s)$, $s \downarrow 0$, for all $x_0 > 0$ and $\varepsilon > 0$,

$$\lim_{s \downarrow 0} \sup_{0 < x \leq x_0} x^{\gamma_0 + \varepsilon} \left| \frac{\frac{F^{\leftarrow}(1-sx) - F^{\leftarrow}(1-s)}{\tilde{a}(s)} - \frac{x^{-\gamma_0} - 1}{\gamma_0}}{\tilde{\Phi}(s)} - \Psi(x) \right| = 0$$

(Drees, 1998, Lemma 2.1). Combining these two results, we obtain

$$\begin{aligned} \sup_{t \in [1/(2k_n), 1]} t^{\gamma_0 + \varepsilon} \Bigg| & \frac{\frac{F^{\leftarrow}\left(1 - U_{[k_n t]+1,n}\right) - F^{\leftarrow}\left(1 - \frac{k_n}{n}\right)}{\tilde{a}(\frac{k_n}{n})} - \frac{\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right)^{-\gamma_0} - 1}{\gamma_0}}{\tilde{\Phi}(\frac{k_n}{n})} \\ & - \Psi\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right) \Bigg| = o_p(1). \end{aligned} \tag{25}$$

Then we have, for $-1/2 < \gamma_0 < 0$ or $\gamma_0 > 0$,

$$\begin{aligned} \frac{X_{n-[k_n t],n} - X_{n-k_n,n}}{\tilde{a}(\frac{k_n}{n})} &= \frac{F^{\leftarrow}\left(1 - U_{[k_n t]+1,n}\right) - F^{\leftarrow}\left(1 - U_{k_n+1,n}\right)}{\tilde{a}(\frac{k_n}{n})} \\ &= \frac{1}{\gamma_0} \left(\frac{n}{k_n} U_{[k_n t]+1,n}\right)^{-\gamma_0} - \frac{1}{\gamma_0} \left(\frac{n}{k_n} U_{k_n+1,n}\right)^{-\gamma_0} + \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right) \\ & \quad - \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{k_n+1,n}\right) + o_p\Big(t^{-(\gamma_0 + \varepsilon)} \tilde{\Phi}(\frac{k_n}{n})\Big). \end{aligned}$$

Hence

$$
\begin{aligned}
1 &+ \frac{\gamma}{\tilde{\sigma}} \frac{X_{n-[k_n t],n} - X_{n-k_n,n}}{\tilde{a}(\frac{k_n}{n})} \\
&= \left(1 - \left(\frac{n}{k_n} U_{k_n+1,n}\right)^{-\gamma_0}\right) - (\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{1}{\gamma_0}\left(\frac{n}{k_n} U_{k_n+1,n}\right)^{-\gamma_0} \\
&\quad + \frac{\gamma}{\tilde{\sigma}} \frac{1}{\gamma_0}\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right)^{-\gamma_0} + \frac{\gamma}{\tilde{\sigma}} \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right) \\
&\quad - \frac{\gamma}{\tilde{\sigma}} \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{k_n+1,n}\right) + o_p\left(t^{-(\gamma_0+\varepsilon)} k_n^{-1/2}\right) \\
&= I + II + III + IV + V + VI.
\end{aligned}
$$

By (24) $t^{\gamma_0} III$ is bounded away from zero uniformly for $t \in [(2k_n)^{-1}, 1]$. We will show that all the other terms tend to 0 uniformly when multiplied with t_0^{γ}, so that assertion (21) follows with $C_n := \inf_{t \in [(2k_n)^{-1},1]} t^{\gamma_0} III - \varepsilon_n$ for a suitable sequence $\varepsilon_n \downarrow 0$.
By the asymptotic normality of intermediate order statistics, part I is $O_p(k_n^{-1/2})$. Hence $t^{\gamma_0} I = o_p(1)$, which is trivial if $\gamma_0 > 0$; for $-1/2 < \gamma_0 < 0$ note that $t^{\gamma_0} k_n^{-1/2} \le 2^{-\gamma_0} k_n^{-\gamma_0-1/2} \to 0$ as $k_n \to \infty$. By (24) and assumption (20), part II is $O_p(k_n^{-1/2})$ so that by the same arguments as above $t^{\gamma_0} II = o_P(1)$.
Next note that $t^{\gamma_0}\Psi(t) = o(t^{-1/2})$ as $t \downarrow 0$. This combined with (9) and (24) gives that $t^{\gamma_0} IV$ and $t^{\gamma_0} V$ are $o_p(1)$. Finally, $t^{\gamma_0} VI = o_p(1)$, provided one chooses $\varepsilon < 1/2$.
Now consider the case $\gamma_0 = 0$. Since (25) is still valid when $\gamma_0 = 0$, with the obvious changes, we get

$$
\begin{aligned}
&\frac{X_{n-[k_n t],n} - X_{n-k_n,n}}{\tilde{a}(\frac{k_n}{n})} \\
&= -\log\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right) + \log\left(\frac{n}{k_n} U_{k_n+1,n}\right) + \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right) \\
&\quad - \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{k_n+1,n}\right) + o_p\left(t^{-\varepsilon} \tilde{\Phi}(\frac{k_n}{n})\right). \qquad (26)
\end{aligned}
$$

Hence

$$
\begin{aligned}
1 &+ \frac{\gamma}{\tilde{\sigma}} \frac{X_{n-[k_n t],n} - X_{n-k_n,n}}{\tilde{a}(\frac{k_n}{n})} \\
&= 1 - \frac{\gamma}{\tilde{\sigma}} \log t - \frac{\gamma}{\tilde{\sigma}} \log\left(\frac{n}{k_n t} U_{[k_n t]+1,n}\right) + \frac{\gamma}{\tilde{\sigma}} \log\left(\frac{n}{k_n} U_{k_n+1,n}\right) \\
&\quad + \frac{\gamma}{\tilde{\sigma}} \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{[k_n t]+1,n}\right) - \frac{\gamma}{\tilde{\sigma}} \tilde{\Phi}(\frac{k_n}{n}) \Psi\left(\frac{n}{k_n} U_{k_n+1,n}\right) + o_p(\frac{\gamma}{\tilde{\sigma}} t^{-\varepsilon} k_n^{-1/2}).
\end{aligned}
$$

Hence by (24) and assumptions (9) and (20), all the terms but the 1 in the last equality tend to 0 in probability uniformly for $t \in [1/(2k_n), 1]$ so that (22) is obvious.

Finally, to verify (23) just note that for $t = 1/(2k_n)$ the expression (26) is of the order $O_p(\log k_n)$, provided $0 < \varepsilon < 1/2$. Since $X_{n-[k_n t],n} \le X_{n,n}$ for all $t \in [0,1]$ the assertion (23) follows. $\square$

Proposition 3.1. *Assume conditions* (7) *and* (9). *Any solution* $(\gamma, \tilde\sigma)$ *of* (19) *satisfying* (20) *and* $\log\tilde\sigma = O_P(1)$ *admits the approximation*

$$\begin{aligned} &k_n^{1/2}(\gamma - \gamma_0) - \frac{(\gamma_0+1)^2}{\gamma_0}\int_0^1 \left(t^{\gamma_0} - (2\gamma_0+1)t^{2\gamma_0}\right) Y_n(t)dt = o_p(1) \\ &k_n^{1/2}(\tilde\sigma - 1) - \frac{\gamma_0+1}{\gamma_0}\int_0^1 ((\gamma_0+1)(2\gamma_0+1)t^{2\gamma_0} - t^{\gamma_0})Y_n(t)dt = o_p(1), \end{aligned} \tag{27}$$

as $n \to \infty$. *For* $\gamma_0 = 0$ *these equations should be interpreted as their limits for* $\gamma_0 \to 0$, *i.e.*

$$\begin{aligned} &k_n^{1/2}\gamma + \int_0^1 (2 + \log t)Y_n(t)dt = o_p(1) \\ &k_n^{1/2}(\tilde\sigma - 1) - \int_0^1 (3 + \log t)Y_n(t)dt = o_p(1). \end{aligned} \tag{28}$$

Conversely, there exists a solution of (19) *which satisfies* (27) *respectively* (28).

Remark 3.1. For $\gamma_0 \neq 0$ the condition on $\log\tilde\sigma$ is not needed.

Proof. We consider the cases $\gamma_0 > 0$, $-1/2 < \gamma_0 < 0$ and $\gamma_0 = 0$ separately.
Case $\gamma_0 > 0$. In view of assumption (20), we may assume $\gamma \neq 0$. Hence system (19) can be simplified to

$$\begin{cases} \displaystyle\int_0^1 \log\left(1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t) - Q_n(1)}{\tilde a(\frac{k_n}{n})}\right) dt = \gamma \\ \displaystyle\int_0^1 \frac{1}{1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}} dt = \frac{1}{\gamma+1}. \end{cases} \tag{29}$$

Next we will find expansions for the left hand side of both equations.
Rewrite the first one as

$$\begin{aligned} &\int_0^{(2k_n)^{-1}} \log\left(1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t) - Q_n(1)}{\tilde a(\frac{k_n}{n})}\right) dt + \int_{(2k_n)^{-1}}^1 \log(t^{-\gamma_0})\, dt \\ &\quad + \int_{(2k_n)^{-1}}^1 \log\left(t^{\gamma_0}\left(1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t) - Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)\right) dt \\ &\quad = I_1 + \gamma_0(1 - O(k_n^{-1}\log k_n)) + I_2. \end{aligned}$$

First we prove that I_1 is negligible. Since $t \mapsto Q_n(t)$ is constant when $t \in [0, (2k_n)^{-1}]$, Lemma 3.2 implies that, with probability tending to 1,

$$1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t) - Q_n(1)}{\tilde a(\frac{k_n}{n})} = 1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(\frac{1}{2k_n}) - Q_n(1)}{\tilde a(\frac{k_n}{n})} \ge (2k_n)^{\gamma_0} C_n \tag{30}$$

for all $t \in [0,(2k_n)^{-1}]$ with C_n stochastically bounded away from 0, so that $-I_1 \le (2k_n)^{-1} O_P(\log k_n)$. On the other hand from (16), (18) and (20),

$$
\begin{aligned}
1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(\frac{1}{2k_n})-Q_n(1)}{\tilde a(\frac{k_n}{n})} &= 1+\big(\gamma_0+O_P(k_n^{-1/2})\big)\Big(\frac{(2k_n)^{\gamma_0}-1}{\gamma_0}+O_P(k_n^{\gamma_0+\varepsilon})\Big) \\
&= O_P(k_n^{\gamma_0+\varepsilon}).
\end{aligned}
$$

Hence it follows that $I_1 = o_p(k_n^{-1/2})$.
Next we turn to the main term I_2. We will apply the inequality $0 \le x - \log(1+x) \le x^2/(2(1\wedge(1+x))$, valid for all $x>-1$, to

$$
x = t^{\gamma_0}\left(1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)-1 = \Big(\frac{\gamma}{\tilde\sigma}-\gamma_0\Big)\frac{1-t^{\gamma_0}}{\gamma_0}+\frac{\gamma}{\tilde\sigma}k_n^{-1/2}t^{\gamma_0}Y_n(t). \tag{31}
$$

Then, from Lemma 3.2 it follows that $0 < 1/(1\wedge(1+x)) \le 1 \vee 1/C_n = O_P(1)$ with probability tending to one. Moreover note that relation (18) implies

$$
\int_0^{(2k_n)^{-1}} t^{\gamma_0}Y_n(t)dt = O_P\left(\int_0^{(2k_n)^{-1}} t^{-1/2-\varepsilon}dt\right) = O_p((2k_n)^{-1/2+\varepsilon}) = o_p(1),
$$

for $\varepsilon \in (0,1/2)$. Hence from (16) and (18), as $n\to\infty$,

$$
\begin{aligned}
I_2 &= \int_{(2k_n)^{-1}}^1 t^{\gamma_0}\Big(1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\Big)-1\,dt \\
&\quad +O_P\Big(\int_{(2k_n)^{-1}}^1\Big(t^{\gamma_0}\Big(1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\Big)-1\Big)^2dt\Big) \\
&= \int_{(2k_n)^{-1}}^1\Big((\frac{\gamma}{\tilde\sigma}-\gamma_0)\frac{1-t^{\gamma_0}}{\gamma_0}+\frac{\gamma}{\tilde\sigma}k_n^{-1/2}t^{\gamma_0}Y_n(t)\Big)\,dt \\
&\quad +O_p\left(\int_{(2k_n)^{-1}}^1\Big((\frac{\gamma}{\tilde\sigma}-\gamma_0)\frac{1-t^{\gamma_0}}{\gamma_0}+\frac{\gamma}{\tilde\sigma}k_n^{-1/2}t^{\gamma_0}Y_n(t)\Big)^2\,dt\right) \\
&= \Big((\frac{\gamma}{\tilde\sigma}-\gamma_0)\frac{1}{(\gamma_0+1)}+O_p(k_n^{-1/2}(2k_n)^{-1})\Big) \\
&\quad +\Big(\frac{\gamma}{\tilde\sigma}k_n^{-1/2}\int_0^1 t^{\gamma_0}Y_n(t)\,dt+o_p(k_n^{-1/2})\Big) \\
&\quad +O_p(k_n^{-1}+k_n^{-1}(2k_n)^{2\varepsilon}+k_n^{-1}(2k_n)^{-1/2+\varepsilon}) \\
&= (\frac{\gamma}{\tilde\sigma}-\gamma_0)\frac{1}{(\gamma_0+1)}+\frac{\gamma}{\tilde\sigma}k_n^{-1/2}\int_0^1 t^{\gamma_0}Y_n(t)\,dt+o_p(k_n^{-1/2}),
\end{aligned}
$$

where for the last equality we took $\varepsilon < 1/4$. To sum up, we have proved that

$$
\begin{aligned}
&\int_0^1\log\left(1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)dt \\
&= \gamma_0+(\frac{\gamma}{\tilde\sigma}-\gamma_0)\frac{1}{(\gamma_0+1)}+\frac{\gamma}{\tilde\sigma}k_n^{-1/2}\int_0^1 t^{\gamma_0}Y_n(t)dt+o_p(k_n^{-1/2}).
\end{aligned}
$$

This means that the first equation of (29) is equivalent to

$$\gamma = \gamma_0 + (\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{1}{(\gamma_0+1)} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} \int_0^1 t^{\gamma_0} Y_n(t)dt + o_p(k_n^{-1/2}).$$

Now we deal with the left hand side of the second equation in (29). Applying the equality

$$\frac{1}{1+x} = 1 - x + \frac{x^2}{1+x}$$

valid for $x \neq -1$, to x defined in (31), we get, for $1/(2k_n) \leq t \leq 1$,

$$\begin{aligned}\frac{1}{1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}} &= t^{\gamma_0}\Big[1 - \left(\frac{\gamma}{\tilde{\sigma}} - \gamma_0\right)\frac{1-t^{\gamma_0}}{\gamma_0} - t^{\gamma_0}\frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} Y_n(t) \\ &\quad + \frac{\left(\left(\frac{\gamma}{\tilde{\sigma}} - \gamma_0\right)\frac{1-t^{\gamma_0}}{\gamma_0} + t^{\gamma_0}\frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} Y_n(t)\right)^2}{t^{\gamma_0}\left(1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)}\Big].\end{aligned}$$

Hence the left hand side of the second equation in (29) equals

$$\begin{aligned}&\int_0^{(2k_n)^{-1}} \frac{1}{1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}}\, dt \\ &+ \int_{(2k_n)^{-1}}^1 \left(t^{\gamma_0} - (\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{t^{\gamma_0} - t^{2\gamma_0}}{\gamma_0} - \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{2\gamma_0} Y_n(t)\right) dt \qquad (32)\\ &+ \int_{(2k_n)^{-1}}^1 \frac{\left((\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{1-t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t)\right)^2}{1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}}\, dt.\end{aligned}$$

From (30) it follows easily that the first integral is $o_p(k_n^{-1/2})$. Direct calculations and (18) show that the second integral equals

$$\begin{aligned}&\frac{1}{\gamma_0+1} - (\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{1}{(\gamma_0+1)(2\gamma_0+1)} - \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2}\int_0^1 t^{2\gamma_0} Y_n(t)dt \\ &+ O_p((2k_n)^{-\gamma_0-1} + k_n^{-1/2}(2k_n)^{-\gamma_0-1} + k_n^{-1/2}(2k_n)^{-\gamma_0-1/2+\varepsilon})\end{aligned}$$

Here for $\varepsilon < 1/2$ the O_p-term is $o_p(k_n^{-1/2})$. By Lemma 3.2, the last integral of (32) is bounded by

$$\begin{aligned}&O_p\left(\int_{(2k_n)^{-1}}^1 t^{\gamma_0}\left((\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{1-t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t)\right)^2 dt\right) \\ =\ & O_p(k_n^{-1} + k_n^{-1}(1+(2k_n)^{-\gamma_0+2\varepsilon}) + k_n^{-1}(2k_n)^{-\gamma_0-1/2+\varepsilon}) = o_p(k_n^{-1/2}),\end{aligned}$$

if $\varepsilon < 1/4 + \gamma_0/2$. Therefore we have proved

$$\begin{aligned}\int_0^1 \frac{1}{1+\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}} dt &= \frac{1}{\gamma_0+1} - (\frac{\gamma}{\tilde{\sigma}} - \gamma_0)\frac{1}{(\gamma_0+1)(2\gamma_0+1)} \\ &\quad - \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2}\int_0^1 t^{2\gamma_0} Y_n(t)dt + o_p(k_n^{-1/2}).\end{aligned}$$

Hence, under the given conditions, system (29) is equivalent to

$$\begin{cases} \gamma_0 + (\frac{\gamma}{\tilde\sigma} - \gamma_0)\frac{1}{\gamma_0+1} + \frac{\gamma}{\tilde\sigma}k_n^{-1/2}\int_0^1 t^{\gamma_0}Y_n(t)dt + o_p(k_n^{-1/2}) = \gamma \\ \frac{1}{\gamma_0+1} - (\frac{\gamma}{\tilde\sigma} - \gamma_0)\frac{1}{(\gamma_0+1)(2\gamma_0+1)} - \frac{\gamma}{\tilde\sigma}k_n^{-1/2}\int_0^1 t^{2\gamma_0}Y_n(t)dt + o_p(k_n^{-1/2}) = \frac{1}{\gamma+1}. \end{cases} \tag{33}$$

Next we prove that (33) implies (27). First note that, in view of (18) and (20), (33) implies

$$\begin{cases} \gamma_0 + (\frac{\gamma}{\tilde\sigma} - \gamma_0)\frac{1}{\gamma_0+1} + \gamma_0 k_n^{-1/2}\int_0^1 t^{\gamma_0}Y_n(t)dt + o_p(k_n^{-1/2}) = \gamma \\ \frac{1}{\gamma_0+1} - (\frac{\gamma}{\tilde\sigma} - \gamma_0)\frac{1}{(\gamma_0+1)(2\gamma_0+1)} - \gamma_0 k_n^{-1/2}\int_0^1 t^{2\gamma_0}Y_n(t)dt + o_p(k_n^{-1/2}) = \frac{1}{\gamma+1}. \end{cases} \tag{34}$$

The first equation and (20) show that $|\gamma-\gamma_0| = O_p(k_n^{-1/2})$, hence $|\gamma-\gamma_0|^2 = o_p(k_n^{-1/2})$. Therefore $1/(\gamma_0+1) - 1/(\gamma+1) = (\gamma-\gamma_0)/(\gamma_0+1)^2 + o(k_n^{-1/2})$ and so (34) implies

$$\begin{cases} \gamma - \gamma_0 - (\frac{\gamma}{\tilde\sigma} - \gamma_0)\frac{1}{\gamma_0+1} - k_n^{-1/2}\gamma_0\int_0^1 t^{\gamma_0}Y_n(t)dt + o_p(k_n^{-1/2}) = 0 \\ \frac{\gamma-\gamma_0}{(\gamma_0+1)^2} - (\frac{\gamma}{\tilde\sigma} - \gamma_0)\frac{1}{(\gamma_0+1)(2\gamma_0+1)} - k_n^{-1/2}\gamma_0\int_0^1 t^{2\gamma_0}Y_n(t)dt + o_p(k_n^{-1/2}) = 0. \end{cases}$$

Now straightforward calculations show that a solution of this linear system in $\gamma-\gamma_0$ and $\gamma/\tilde\sigma - \gamma_0$ satisfies (27).
Since conversely a solution of type (27) obviously satisfies the condition (20), it is easily seen that it also solves (33) and thus (19).
Case $-1/2 < \gamma_0 < 0$. Again, in this case system (19) simplifies to (29). Rewrite the left hand side of the first equation as

$$\begin{aligned} &\int_0^{s_n} \log\left(1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)dt + \int_{s_n}^1 \log(t^{-\gamma_0})\,dt \\ &+ \int_{s_n}^1 \log\left(t^{\gamma_0}\left(1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)\right)dt \\ &= J_1 + (\gamma_0 + O(s_n|\log s_n|)) + J_2 \end{aligned}$$

and choose $s_n = k_n^{-\delta}$, with $\delta \in (1/2, (4\varepsilon)^{-1})$ for some $\varepsilon \in (0, 1/2)$.
Now we prove that J_1 is negligible. Note that since $t \mapsto Q_n(t)$ is constant when $t \in [0, (2k_n)^{-1}]$, (21) is trivially extended to $t \in [0,1]$ when $\gamma_0 < 0$. By definition $Q_n(t) - Q_n(1) \geq 0$, for all $t \in [0,1]$ and $\tilde a(k_n/n) > 0$. Since by (20) $P\{\gamma < 0\} \to 1$, Lemma 3.2 implies

$$P\left(\left|\log\left(1 + \frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)\right| \leq |\log(C_n t^{-\gamma_0})|, \quad t \in [0,1]\right) \to 1,$$

as $n \to \infty$, so that $\int_0^{s_n} |\log(C_n t^{-\gamma_0})|\,dt = O_P(s_n|\log s_n|) = o_p(k_n^{-1/2})$ gives $J_1 = o_p(k_n^{-1/2})$.
Next we approximate J_2. Check that $0 \leq x - \log(1+x) \leq x^2/[2(1 \wedge (1+x))]$ holds for all $x > -1$. Hence, in view of (16) and Lemma 3.2, choosing

$x = t^{\gamma_0}[1 + (\gamma/\tilde{\sigma})(Q_n(t) - Q_n(1))/a(k_n/n)]$, we obtain

$$
\begin{aligned}
J_2 = &\int_{s_n}^1 \left((\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1 - t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t) \right) dt \\
&+ O_p \left(\int_{s_n}^1 \left[(\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1 - t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t) \right]^2 dt \right) \\
= &\left((\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1}{\gamma_0 + 1} + O_p(k_n^{-1/2} s_n^{\gamma_0 + 1}) \right) \\
&+ \left(\frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} \int_0^1 t^{\gamma_0} Y_n(t) dt + O_p(k_n^{-1/2} \int_0^{s_n} t^{-1/2-\varepsilon} dt) \right) \\
&+ O_p \left(\int_{s_n}^1 \left[(\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1 - t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t) \right]^2 dt \right)
\end{aligned}
$$

and from the choice of s_n, (20) and $\gamma_0 \in (-1/2, 0)$ it follows that the O_p-terms are $o_p(k_n^{-1/2})$. Hence we proved that

$$
\begin{aligned}
&\int_0^1 \log \left(1 + \frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})} \right) dt \\
&= \gamma_0 + (\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1}{\gamma_0 + 1} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} \int_0^1 t^{\gamma_0} Y_n(t) dt + o_p(k_n^{-1/2}).
\end{aligned}
$$

Now we turn to the second equation in (29). Use a similar decomposition as in the case $\gamma > 0$ of the left hand side:

$$
\begin{aligned}
&\int_0^{s_n} \frac{1}{1 + \frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}} dt + \int_{s_n}^1 \left(t^{\gamma_0} - (\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{t^{\gamma_0} - t^{2\gamma_0}}{\gamma_0} - \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{2\gamma_0} Y_n(t) \right) dt \\
&\quad + \int_{s_n}^1 \frac{\left((\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1 - t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t) \right)^2}{1 + \frac{\gamma}{\tilde{\sigma}} \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}} dt \\
&=: \; K_1 + K_2 + K_3,
\end{aligned}
$$

with $s_n = k_n^{-\delta}$ for some $\delta \in ((2\gamma_0 + 2)^{-1}, (4\varepsilon - 2\gamma_0)^{-1})$ and $\varepsilon \in (0, \gamma_0 + 1/2)$. Then by Lemma 3.2, $K_1 = O_p(s_n^{\gamma_0+1}) = o_p(k_n^{-1/2})$. Moreover

$$
\begin{aligned}
K_2 \;=\; &\frac{1}{\gamma_0 + 1} - (\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1}{(\gamma_0 + 1)(2\gamma_0 + 1)} - \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} \int_0^1 t^{2\gamma_0} Y_n(t) dt \\
&+ O_p(s_n^{\gamma_0+1} + k_n^{-1/2} s_n^{2\gamma_0+1} + k_n^{-1/2} s_n^{\gamma_0+1/2-\varepsilon})
\end{aligned}
$$

and, by the choice of δ, we have that the O_p-term is $o_p(k_n^{-1/2})$. Finally from Lemma 3.2 and the definition of s_n,

$$
\begin{aligned}
K_3 \;&=\; O_p \left(\int_{s_n}^1 t^{\gamma_0} \left[(\frac{\gamma}{\tilde{\sigma}} - \gamma_0) \frac{1 - t^{\gamma_0}}{\gamma_0} + \frac{\gamma}{\tilde{\sigma}} k_n^{-1/2} t^{\gamma_0} Y_n(t) \right]^2 dt \right) \\
&=\; O_p \left(k_n^{-1} (s_n^{3\gamma_0+1} \vee 1) + k_n^{-1} (s_n^{\gamma_0 - 2\varepsilon} + 1) + k_n^{-1} (s_n^{2\gamma_0 + 1/2 - \varepsilon} + 1) \right) \\
&=\; o_p(k_n^{-1/2}).
\end{aligned}
$$

Hence the proof can be concluded by the same arguments as in the case $\gamma_0 > 0$.
Case $\gamma_0 = 0$. In this case we use equations (19). Apply (twice) the equality $1/(1+x) = 1 - x + x^2/(1+x)$, the inequality $|x - \log(1+x) - x^2/2 + x^3/3| \leq x^4/[4(1 \wedge (1+x)^4)]$, valid for all $x > -1$, to $x = (\gamma/\tilde{\sigma})(Q_n(t) - Q_n(1))/\tilde{a}(k_n/n)$, and use (22) in Lemma 3.2, to obtain for the left hand side of the first equation

$$\frac{1}{\gamma^2}\int_0^1 \log\left(1 + \frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right) - (1+\gamma)\frac{\frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}}{1 + \frac{\gamma}{\tilde{\sigma}}\frac{Q_n(t)-Q_n(1)}{\tilde{a}(\frac{k_n}{n})}}\,dt$$

$$= \int_0^1 -\frac{1}{\tilde{\sigma}}\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\,dt + \int_0^1 \left(\frac{1}{2} + \gamma\right)\frac{1}{\tilde{\sigma}^2}\left(\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)^2 dt$$

$$- \int_0^1 \left(\frac{2}{3} + \gamma\right)\frac{\gamma}{\tilde{\sigma}^3}\left(\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)^3 dt$$

$$+ O_p\left(\int_0^1 \frac{\gamma^2}{\tilde{\sigma}^4}\left[\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right]^4 + \frac{\gamma^3}{\tilde{\sigma}^5}\left[\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right]^5 dt\right). \tag{35}$$

By (16) the first integral in the right hand side of the last equation equals $-\tilde{\sigma}^{-1} - \tilde{\sigma}^{-1}k_n^{-1/2}\int_0^1 Y_n(t)dt + o_p(k_n^{-1/2})$. For the second integral in the right hand side of (35) consider

$$\int_0^{s_n} \left(\frac{1}{2} + \gamma\right)\frac{1}{\tilde{\sigma}^2}\left(\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)^2 dt$$
$$+ \int_{s_n}^1 \left(\frac{1}{2} + \gamma\right)\frac{1}{\tilde{\sigma}^2}\left(\frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\right)^2 dt,$$

with $s_n = k_n^{-\delta}$, $\delta \in (1/2, (4\varepsilon)^{-1})$, $\varepsilon \in (0, 1/2)$. Then the first of these last two integrals is $o_p(k_n^{-1/2})$ by (23). In view of (16), (18)and $|\gamma/\tilde{\sigma}| = O_p(k_n^{-1/2})$, the second integral equals

$$\frac{1+2\gamma}{\tilde{\sigma}^2} - \frac{1}{\tilde{\sigma}^2}k_n^{-1/2}\int_0^1 (\log t)Y_n(t)dt + o_p(k_n^{-1/2}).$$

Using a similar reasoning, but with $\delta \in (1/2, 3(4\varepsilon)^{-1} \wedge 4(1+6\varepsilon)^{-1})$, $\varepsilon \in (0, 1/2)$, the third integral of (35) equals $-4\gamma\tilde{\sigma}^{-3} + o_p(k_n^{-1/2})$. Finally the O_p-term of (35) is clearly $o_p(k_n^{-1/2})$ by (23).
Hence we have that (35) equals

$$-\frac{1}{\tilde{\sigma}} + \frac{1+2\gamma}{\tilde{\sigma}^2} - \frac{4\gamma}{\tilde{\sigma}^3} - \frac{1}{\tilde{\sigma}}k_n^{-1/2}\int_0^1 Y_n(t)dt - \frac{1}{\tilde{\sigma}^2}k_n^{-1/2}\int_0^1 (\log t)Y_n(t)dt + o_p(k_n^{-1/2}).$$

To deal with the left hand side of the second equation, use again the aforementioned equality for $1/(1+x)$ and (22) in Lemma 3.2 to get

$$\begin{aligned}
&\frac{1+\gamma}{\gamma}\left[\int_0^1 \frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}dt - \int_0^1\left(\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)^2 dt\right.\\
&\qquad\left. +O_p\left(\int_0^1\left[\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right]^3 dt\right)\right]\\
&= (1+\gamma)\left[\frac{1}{\tilde\sigma}+\frac{1}{\tilde\sigma}k_n^{-1/2}\int_0^1 Y_n(t)dt - \int_0^{s_n}\gamma\left(\frac{1}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)^2 dt\right.\\
&\qquad\left. -\int_{s_n}^1\gamma\left(\frac{1}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}\right)^2 dt + O_p\left(k_n^{-1}(\log k_n)^3\right)\right],
\end{aligned}$$

where for the O_p-term we used (23) and $\gamma = O_p(k_n^{-1/2})$. Next we consider the second and third integral in the last equality, L_1 and L_2 say. As for L_1, it follows from (23) that

$$L_1 = O_p\left(s_n\gamma(\log k_n)^2\right) = o_p(k_n^{-1/2})$$

if $s_n = k_n^{-\delta}$, $\delta\in(0,1)$. As for L_2, from (18) with $\varepsilon\in(0,1/2)$, we get

$$L_2 = -2\frac{\gamma}{\tilde\sigma^2} + O_p\left(k_n^{-3/2}s_n^{-2\varepsilon}+k_n^{-1}\right) = -2\frac{\gamma}{\tilde\sigma^2}+o_p(k_n^{-1/2}).$$

Hence we proved that

$$\int_0^1\left(\frac{1}{\gamma}+1\right)\frac{\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}}{1+\frac{\gamma}{\tilde\sigma}\frac{Q_n(t)-Q_n(1)}{\tilde a(\frac{k_n}{n})}}dt = \frac{1+\gamma}{\tilde\sigma}-\frac{2\gamma}{\tilde\sigma^2}+\frac{1}{\tilde\sigma}k_n^{-1/2}\int_0^1 Y_n(t)dt + o_p(k_n^{-1/2}).$$

Therefore, under the given conditions, a solution of (19) must satisfy

$$\begin{cases}
(1-\tilde\sigma)+2\gamma-\frac{4\gamma}{\tilde\sigma}-\tilde\sigma k_n^{-1/2}\int_0^1 Y_n(t)dt - k_n^{-1/2}\int_0^1\log tY_n(t)dt + o_p(k_n^{-1/2}) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = 0\\
(1-\tilde\sigma)+\gamma-\frac{2\gamma}{\tilde\sigma}+k_n^{-1/2}\int_0^1 Y_n(t)dt + o_p(k_n^{-1/2}) = 0.
\end{cases}$$

Next note that the first equation implies $\tilde\sigma = 1+O_p(k_n^{-1/2})$, and so $\gamma/\tilde\sigma = \gamma + o_p(k_n^{-1/2})$. Simplifying the above equations, we arrive at (28).
The converse assertion is proved as in the case $\gamma_0>0$. □

Proof of Theorem 2.1. Because of $\tilde a(k_n/n)/a(k_n/n)-1 = o(k_n^{-1/2})$, Proposition 3.1 shows that, under the conditions (7) and (9), any solution $(\hat\gamma_n^*,\hat\sigma_n^*)$ of the likelihood equations such that $k_n^{1/2}|\hat\gamma_n^*-\gamma_0|$ and $k_n^{1/2}|\hat\sigma_n^*/a(k_n/n)-1|$ are stochastically bounded, must satisfy (27)–(28) and hence, in view of Lemma 3.1 and $\Phi(k_n/n)\sim\tilde\Phi(k_n/n)$, also (10)–(13).
Conversely, according to Proposition 3.1, there exists a solution of (19) satisfying (27) (respectively (28)). This solution corresponds to a solution of the likelihood equations satisfying (10)–(13). □

Proof of Corollary 2.1. According to Theorem 2.1, the components of the left hand side of (15) minus deterministic bias terms converge to certain integrals of a Gaussian process, that is, to normal random variables. If $k_n^{1/2}\Phi(k_n/n) \to \lambda$, the bias term of $k_n^{1/2}(\hat{\gamma}_n - \gamma_0)$ tends to $\lambda((\gamma_0+1)^2/\gamma_0)\int_0^1 (t^{\gamma_0} - (2\gamma_0+1)t^{2\gamma_0})\Psi(t)\,dt$. Using (8) the result follows by simple calculations. Similarly the asymptotic bias of the second component can be derived.
To calculate the variance of the limiting normal random variable corresponding to $k_n^{1/2}(\hat{\gamma}_n - \gamma_0)$, let $X(t) = ((\gamma_0+1)^2/\gamma_0)\ (t^{\gamma_0} - (2\gamma_0+1)t^{2\gamma_0})\big(W(1) - t^{-(\gamma_0+1)}W(t)\big)$. Then straightforward calculations show that var $(\int_0^1 X(t)dt) = \int_0^1\int_0^1 E[X(s)X(t)]\,ds\,dt = (\gamma_0+1)^2$.
Likewise, to obtain the asymptotic covariance of $k_n^{1/2}(\hat{\sigma}_n/a(k_n/n) - 1)$ with $k_n^{1/2}(\hat{\gamma}_n - \gamma_0)$, let $Y(t) = ((\gamma_0+1)/\gamma_0)\,((\gamma_0+1)(2\gamma_0+1)t^{2\gamma_0} - t^{\gamma_0})\big(W(1) - t^{-(\gamma_0+1)}W(t)\big)$. Then $\mathrm{cov}(\int_0^1 X(s)\,ds, \int_0^1 Y(t)\,dt) = \int_0^1\int_0^1 E[X(s)Y(t)]\,ds\,dt = -(1+\gamma_0)$. The limiting variance of the scale estimator is obtained similarly. □

Proof of Theorem 2.2. Integration of the various terms of (17) yields for $\gamma_0 = 0$

$$\begin{aligned} & k_n^{1/2}\left(\frac{m_n^{(1)}}{\tilde{a}(\frac{k_n}{n})} - 1\right) \\ = \ & k_n^{1/2}\left(\int_0^1 \frac{Q_n(t) - Q_n(1)}{\tilde{a}(\frac{k_n}{n})}\,dt + \int_0^1 \log t\,dt\right) \\ = \ & \int_0^1 \big(W_n(1) - t^{-1}W_n(t)\big)\ dt + k_n^{1/2}\tilde{\Phi}(\frac{k_n}{n})\int_0^1 \Psi(t)\,dt + o_p(1). \end{aligned}$$

Similarly we obtain

$$\begin{aligned} k_n^{1/2}\left(\frac{m_n^{(2)}}{(\tilde{a}(\frac{k_n}{n}))^2} - 2\right) = & -\int_0^1 2\log t\,\big(W_n(1) - t^{-1}W_n(t)\big)\ dt \\ & + k_n^{1/2}\Phi(\frac{k_n}{n})\int_0^1 2\log t\Psi(t)\,dt + o_P(1). \end{aligned}$$

Now, using Taylor expansions, straightforward calculations give

$$\begin{aligned} k_n^{1/2}\hat{\gamma}_* = & -\int_0^1 (2+\log t)\,\big(W_n(1) - t^{-1}W_n(t)\big)\ dt \\ & + k_n^{1/2}\Phi(\frac{k_n}{n})\int_0^1 (2+\log t)\Psi(t)\,dt + o_P(1), \end{aligned}$$

and hence by (12)

$$k_n^{1/2}\,(\hat{\gamma}_* - \hat{\gamma}_{MLE}) \to_p 0.$$

The proof of the second statement is similar. □

Proof of Remark 2.4. Under the stated conditions the following analogue of (17) holds:

$$
\begin{aligned}
& k_n^{1/2}\left(\frac{\log Q_n(t)-\log Q_n(1)}{\tilde a(\frac{k_n}{n})/F^{\leftarrow}(1-\frac{k_n}{n})}+\log t\right) \\
= \;& W_n(1)-t^{-1}W_n(t)-k_n^{1/2}\Phi^*(\frac{k_n}{n})\frac{\log^2 t}{2}+o_p(1)t^{-1/2-\varepsilon},
\end{aligned}
$$

for some function Φ^* such that

$$
\Phi^*(\frac{k_n}{n}) \sim \frac{\tilde a(\frac{k_n}{n})}{F^{\leftarrow}(1-\frac{k_n}{n})} = O(k_n^{-1/2}), \tag{36}
$$

see Draisma et al. (1999), Appendix. Now the results by de Haan and Stadtmüller (1996) imply that $\tilde a(k_n/n)$ tends to a positive constant, while $F^{\leftarrow}(1-k_n/n)$ behaves like a multiple of $\log(n/k_n)$. Hence the bias term is asymptotically negligible if $k_n = o(\log^2 n)$, and the assertion can be concluded by the same reasoning as in the proof of Theorem 2.2. □

Acknowledgement: Holger Drees was supported by Netherlands Organization for Scientific Research through the Netherlands Mathematical Research Foundation and by the Heisenberg program of the DFG. Ana Ferreira's research was partially supported by POCTI/FCT/FEDER.

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H. N. (1992) *A first course in order statistics.* John Wiley & Sons Inc.

[2] Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probab.* 2, 792–804.

[3] Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* 17, 1833–1855.

[4] Draisma, G., de Haan, L., Peng, L. and Pereira, T.T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme-value index. *Extremes* 2, 367–404.

[5] Drees, H. (1998). On smooth statistical tail functionals. *Scand. J. Statist.* 25, 187–210.

[6] Gnedenko, B.V. (1943). Sur la distribution limite du terme du maximum d'une série aléatoire. *Ann. Math.* 44, 423–453.

[7] Grimshaw, S.D. (1993). Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics* 35, 185–191.

[8] de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Australian Math. Soc. (Series A)* 61, 381–395.

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

[10] Pickands, J. III (1975). Statistical inference using extreme order statistics. *Ann. Statist.* 3, 119–131.

[11] Shorack, G. and Wellner, J. (1986). *Empirical processes with applications to statistics.* Wiley, New York.

[12] Smith, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.* 15, 1174–1207.