Ann. Inst. Statist. Math.
Vol. 47, No. 4, 693–717 (1995)

LAN OF EXTREME ORDER STATISTICS

MICHAEL FALK

*Mathematisch-Geographische Fakultät, Katholische Universität Eichstätt,
85071 Eichstätt, Germany*

(Received March 10, 1994; revised December 16, 1994)

Abstract. Consider an iid sample $Z_1, \ldots, Z_n$ with common distribution function F on the real line, whose upper tail belongs to a parametric family $\{F_\beta : \beta \in \Theta\}$. We establish local asymptotic normality (LAN) of the loglikelihood process pertaining to the vector $(Z_{n-i+1:n})_{i=1}^k$ of the upper $k = k(n) \to_{n\to\infty} \infty$ order statistics in the sample, if the family $\{F_\beta : \beta \in \Theta\}$ is in a neighborhood of the family of generalized Pareto distributions. It turns out that, except in one particular location case, the kth-largest order statistic $Z_{n-k+1:n}$ is the central sequence generating LAN. This implies that $Z_{n-k+1:n}$ is asymptotically sufficient and that asymptotically optimal tests for the underlying parameter β can be based on the single order statistic $Z_{n-k+1:n}$. The rate at which $Z_{n-k+1:n}$ becomes asymptotically sufficient is however quite poor.

Key words and phrases: Extreme order statistics, local asymptotic normality, central sequence, generalized Pareto distributions, asymptotic sufficiency, optimal tests.

1. Introduction

Let $Z_1, \ldots, Z_n$ be independent copies of a random variable (rv) Z on the real line with distribution function (df) F. We suppose that the upper tail of F belongs to some parametric family, that is, we assume that

$$\text{(M)} \qquad F(x) = F_\beta(x), \qquad x \geq x_0(\beta),$$

where $\mathcal{F} := \{F_\beta : \beta \in \Theta\}$ is a parametric family of dfs and the point $x_0(\beta)$ is *unknown*.

Such a model for the upper tail of the underlying df is usually indispensable if one is interested in the statistical analysis of extreme quantities of F, which are outside the range of the given data $Z_1, \ldots, Z_n$. Extreme quantiles $F^{-1}(1-q) = \inf\{t \in \mathbb{R} : F(t) \geq 1-q\}$ with q close to zero, are for example needed if one wants to know that height of a dike which is exceeded by the annual maximum flood Z with probability not greater than q (cf. Dekkers and de Haan (1989)).

It is clear from the model (M) that a statistical analysis of the underlying parameter β, which usually has to precede any further investigation, can reasonably be based only on large observations among the sample $Z_1,\dots,Z_n$. A reasonable way is to base an analysis of β on the vector $(Z_{n-i+1:n})_{i=1}^k$ of the k largest order statistics in the sample, where $Z_{1:n}\le\dots\le Z_{n:n}$ denote the ordered values of $Z_1,\dots,Z_n$.

We will utilize in the present paper the order statistics approach for the investigation of the testing problem

$$\beta=\beta_0 \quad \text{against a sequence} \quad \beta=\beta_n$$

of suitable (contiguous) alternatives with a particular underlying family $\mathcal{F}$.

While extreme value statistic has been focussing on the *estimation* of the parameter β in the model (M) (see for example Smith (1987), Dekkers and de Haan (1989), Chapter 9 of Reiss (1989) and the literature cited therein), the testing problem has only recently received increasing attention (Castillo *et al.* (1989), Hasofer and Wang (1992), Falk (1992) among others).

In particular the powerful theory of *local asymptotic normality* (LAN) of *statistical experiments*, developed by LeCam (LeCam (1960, 1986), LeCam and Yang (1990), Strasser (1985)), has been applied to extreme value problems by now only in a few papers (Falk (1992), Marohn (1991, 1994*a*, 1994*b*), Janssen and Marohn (1994), Wei (1992), related papers are Janssen and Reiss (1988) and Marohn (1995)). By this approach the derivation of asymptotically optimal level α tests as well as the computation of their asymptotic power functions is in particular immediate (cf. the discussion after Theorem 2.1).

This was the motivation for the present paper, in which we will prove LAN of the loglikelihood processes pertaining to the vector $(Z_{n-i+1:n})_{i=k}^1$ of the upper order statistics in the underlying model (M) for a particular parametric family $\mathcal{F}$. It will turn out that, except in one particular location problem, the kth-largest order statistic $Z_{n-k+1:n}$, with $k=k(n)\to\infty$ but $k/n\to 0$, is the *central sequence*, generating LAN. This implies that the complete information, which is contained in $(Z_{n-i+1:n})_{i=k}^1$ about the underlying parameter β in the testing problem β_0 against β_n, is asymptotically already contained in the single order statistic $Z_{n-k+1:n}$. In this sense we call $Z_{n-k+1:n}$ asymptotically sufficient.

This result parallels those established in Falk (1992), where it was shown in the peaks over threshold approach to this testing problem, that the *random number* of the exceedances over a sequence of suitable high thresholds carries asymptotically the complete information contained in (the point process of) the exceedances about the underlying parameter β.

Sharp upper bounds for the rate at which $Z_{n-k+1:n}$ becomes asymptotically sufficient, reveal however that this rate is quite poor. The advice to simply drop all order statistics larger than $Z_{n-k+1:n}$ and to base statistical inference in the testing problem above only on $Z_{n-k+1:n}$, can therefore be taken for small up to moderate sample sizes only with a grain of salt.

Next we will introduce the particular parametric class of dfs which we will consider in this paper.

A common condition on $F\in\mathcal{F}$, satisfied by almost any textbook df F, is that F is *in the domain of attraction* of an *extreme value* df G_β, $\beta\in\mathbb{R}$, denoted by

$F \in \mathcal{D}(G_\beta)$. This means that there are constants $a_n > 0$, $b_n \in \mathbb{R}$ such that

$$(Z_{n:n} - b_n)/a_n \underset{\mathcal{D}}{\longrightarrow} G_\beta,$$

where we denote by $\underset{\mathcal{D}}{\longrightarrow}$ convergence in distribution and G_β is defined by

$$G_\beta(x) := \exp(-(1+\beta x)^{-1/\beta}), \quad \text{if} \quad 1 + \beta x > 0.$$

Since $(1+\beta x)^{-1/\beta} \underset{\beta_0}{\longrightarrow} e^{-x}$, $x \in \mathbb{R}$, interpret G_0 as $G_0(x) := \lim_{\beta \to 0} G_\beta(x) = \exp(-e^{-x})$, $x \in \mathbb{R}$. The class G_β, $\beta \in \mathbb{R}$, contains all possible nondegenerate weak limits of $Z_{n:n}$ (Gnedenko (1943), see Galambos ((1987), Chapter 2)).

It was first observed by Balkema and de Haan (1974) and, independently, by Pickands (1975) that $F \in \mathcal{D}(G_\beta)$ if and only if the upper tail of F can be approximated in a suitable sense by the upper tail of a *generalized Pareto distribution* (GPD) H_β, where

$$\begin{aligned} H_\beta(x) &:= 1 + \log(G_\beta(x)), \qquad x \ge 0 \\ &= 1 - (1+\beta x)^{-1/\beta} \begin{cases} \text{for} \quad x \ge 0 & \text{if} \quad \beta \ge 0 \\ \text{for} \quad 0 \le x \le -1/\beta & \text{if} \quad \beta < 0. \end{cases} \end{aligned}$$

Interpret H_0 as $H_0(x) = \lim_{\beta \to 0} H_\beta(x) = 1 - \exp(-x)$, $x \ge 0$. The family of GPDs is already a rather rich one: With $\beta > 0$ we obtain the usual Pareto family, H_{-1} is the uniform distribution on (0,1) and H_0 the standard exponential distribution.

In view of the results by Balkema and de Haan (1974) and Pickands (1975), the family $\{H_\beta : \beta \in \mathbb{R}\}$ is a natural candidate for the class $\mathcal{F} = \{F_\beta : \beta \in \mathbb{R}\}$. This would however exclude the extreme value distributions G_β themselves, which have been typically assumed as underlying dfs in extreme value statistics since the book by Gumbel (1958).

We will therefore consider F_β in certain neighborhoods of GPDs. Denote by $\omega(F) := \sup\{t \in \mathbb{R} : F(t) < 1\} \in (-\infty, \infty]$ the right endpoint of the support of a df F and by h_β the density of the GPD H_β. Then the df F is in a *δ-neighborhood* of a GPD for some $\delta > 0$, iff $\omega(F) = \omega(H_\beta)$ for some $\beta \in \mathbb{R}$ and F has a density f on $(x_0, \omega(F))$ for some $x_0 < \omega(F)$ such that

$$f(x) = h_\beta(x)(1 + O((1 - H_\beta(x))^\delta))$$

as $x \to \omega(F)$. Note that G_β is in a δ-neighborhood of H_β with $\delta = 1$. For a review of the basic role played by δ-neighborhoods of GPDs in extreme value theory we refer to Chapter 2 of Falk *et al.* (1994).

The paper is organized as follows. In Section 2 we establish expansions of the loglikelihood ratios pertaining to the vector $(Z_{n-i+1:n})_{i=1}^k$, which imply LAN of the loglikelihood processes with $Z_{n-k+1:n}$ being the central sequence. This is achieved for an underlying family $\mathcal{F}$ in a δ-neighborhood of the family $\{H_\beta : \beta \in \mathbb{R}\}$ of GPDs with unknown shape parameter β but known scale and location shift. We then establish sharp upper bounds for the rate at which $Z_{n-k+1:n}$ becomes asymptotically sufficient in the sense of asymptotic sufficiency defined by LeCam and Yang ((1990), Section 5.3, Proposition 2). These bounds reveal that $Z_{n-k+1:n}$ becomes asymptotically sufficient only at quite a poor rate.

In Section 3 we add an unknown scale $c_0 > 0$ and location parameter $d_0 \in \mathbb{R}$ to H_{β_0}, and we establish again expansions of the loglikelihood ratios pertaining to $(Z_{n-i+1:n})_{i=k}^{1}$. It turns out that, except in the case $\beta_0 \neq 0$ with unknown location parameter, $Z_{n-k+1:n}$ is again the central sequence generating LAN for suitable sequences of alternatives. It turns further out that if and only if, (β_n, c_n, d_n) lie on a certain hyperplane in $\mathbb{R}^3$, then one cannot distinguish between the alternative (β_n, c_n, d_n) and the hypothesis (β_0, c_0, d_0). This means that the underlying shape parameter β can be hidden by a scale or a location parameter.

Corresponding results for the lower extremes $(Y_{i:n})_{i=1}^{k}$ in a sample $Y_1, \ldots, Y_n$ can be deduced from the equation $(Y_{i:n})_{i=1}^{k} = -(Z_{n-i+1:n})_{i=1}^{k}$, where $Z_i := -Y_i$, $i = 1, \ldots, n$. We index expectation $E_\vartheta(Z)$, distribution $\mathcal{L}_\vartheta(Z)$ etc. of a rv Z by the underlying parameter ϑ. By dP/dQ we denote the density of a probability measure P with respect to a probability measure Q, if it exists.

2. LAN and bounds for the sufficiency

Suppose that the upper tail of F coincides with the upper tail of a GPD i.e., we begin with the model

$$F(x) = H_\beta(x), \qquad x_0(\beta) < x < \omega(H_\beta)$$

for some unknown point $x_0(\beta)$. We will test at first $\beta = 0$ against $\beta = \beta_n$ based on the kth-largest order statistics $(Z_{n-i+1:n})_{i=k}^{1}$, where the sequence $k = k(n)$ satisfies $k \to \infty$, $k/n \to 0$ as $n \to \infty$. This is a crucial condition in order to guarantee that $Z_{n-k+1:n}$ finally exceeds $x_0(\beta)$.

The shape parameter $\beta = 0$ is some kind of change point: If $\beta < 0$, then the right endpoint $\omega(F)$ of F is finite, while in case $\beta > 0$ the right endpoint of F is infinity. The null hypothesis $\beta = 0$ is of a different quality than a hypothetical value $\beta \neq 0$. This is revealed by the observation established in this paper that the rates at which contiguous alternatives β_n converge to β are faster in case $\beta = 0$ than in case $\beta \neq 0$ (see the definition of β_n below and before Theorem 2.3).

Choose $\vartheta \in \mathbb{R}$ and define the alternatives $\beta_n = \beta_n(\vartheta)$ of $\beta = 0$ by

$$\beta_n := 2\vartheta k^{-1/2} / \log^2(n/k).$$

This definition of the alternatives β_n is required for a nondegenerate limit of the loglikelihood ratio in Theorem 2.1. The same applies to subsequent results. We allow the point $x_0(\beta_n)$ to tend to $\omega(H_{\beta_n})$ as n increases, but not too quickly, since we require

$$\limsup_{n\to\infty}(x_0(\beta_n) - \log(n/k)) < 0.$$

Recall that $\omega(H_{\beta_n}) = -1/\beta_n = k^{1/2}\log^2(n/k)/(2|\vartheta|)$ if $\vartheta < 0$ and $\omega(H_{\beta_n}) = \infty$ if $\vartheta \geq 0$.

THEOREM 2.1. (LAN) *Suppose that* $\limsup_{n\to\infty}(x_0(\beta_n) - \log(n/k)) < 0$. *Then we have under the hypothesis* $\beta = 0$

$$\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_0((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1$$
$$= \vartheta k^{1/2}(Z_{n-k+1:n} - \log(n/k)) - \vartheta^2/2 + o_{P_0}(1) \xrightarrow[\mathcal{D}_0]{} N(-\vartheta^2/2, \vartheta^2).$$

The preceding result reveals that the kth-largest order statistics $Z_{n-k+1:n}$ is the *central sequence* for the loglikelihood ratio $\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_0((Z_{n-i+1:n})_{i=k}^1)\}$ that is, all the information about the underlying parameter that is contained in the vector $(Z_{n-i+1:n})_{i=k}^1$ of the upper kth-largest order statistics, is asymptotically already contained in the single order statistic $Z_{n-k+1:n}$. In this sense we may call $Z_{n-k+1:n}$ asymptotically sufficient. (For a precise definition of asymptotic sufficiency we refer to Proposition 2 in Section 5.3 of LeCam and Yang (1990).) Theorem 2.1 parallels Theorem 1.1 in Falk (1992), where it was shown that the *number* of *exceedances* carries asymptotically all the information about the underlying parameter β, which is contained in the *point process* of the exceedances over $t_n := \log(na_n)$ among $Z_1, \ldots, Z_n$.

Theorem 2.1 implies in particular that the finite dimensional marginal distributions of the *loglikelihood process*

$$(X_n(\vartheta))_{\vartheta\in\mathbb{R}} := \left(\log\left\{\frac{d\mathcal{L}_{\beta_n(\vartheta)}((Z_{n-i+1:n})_{i=k}^1)}{d\mathcal{L}_0((Z_{n-i+1:n})_{i=k}^1)}\right\}(Z_{n-i+1:n})_{i=k}^1\right)_{\vartheta\in\mathbb{R}}$$

converge weakly to that of the degenerate Gaussian process

$$(X(\vartheta))_{\vartheta\in\mathbb{R}} := (\vartheta X - \vartheta^2/2)_{\vartheta\in\mathbb{R}},$$

where X is a standard normal rv on the real line, provided $\limsup_{n\to\infty}(x_0(\beta_n) - \log(n/k)) < 0$.

This weak form of convergence of the likelihood processes is one basic definition in LeCam's theory of convergence of statistical experiments. It is already sufficient for example to imply Hájek convolution theorem as well as the Hájek-LeCam asymptotic minimax theorem (cf. Chapter 5 of LeCam and Yang (1990)); details will be given in a subsequent paper. To supply however at least one example of the statistical implications of Theorem 2.1, we demonstrate in the following how it provides asymptotic optimal tests for the hypothesis $\beta = 0$ against $\beta_n(\vartheta)$.

Denote by $u_\alpha = \Phi^{-1}(1-\alpha)$ the $(1-\alpha)$-quantile of the standard normal df Φ. By the Neyman-Pearson lemma and Theorem 2.1, the test

$$\varphi_1(Z_{n-k+1:n}) := 1_{(u_\alpha,\infty)}(k^{1/2}(Z_{n-k+1:n} - \log(n/k)))$$

is an asymptotically optimal level α test based on the kth-largest order statistics for $\beta_n(\vartheta)$ with positive ϑ against $\beta = 0$. As it obviously does not depend on $\vartheta > 0$, this test is even asymptotically optimal *uniformly in* $\vartheta > 0$. This is an example of the advantages of LAN-theory: The loglikelihood ratio for testing $\beta_n(\vartheta)$ against

$\beta_0 = 0$ could be calculated explicitly, but it would depend on ϑ. Moreover, this result remains true, if we require F_β to belong to a δ-neighborhood of a GPD (see (1.5) below), *without* specifying its density *precisely*. The corresponding uniformly asymptotic optimal test for $\beta_n(\vartheta)$ with negative ϑ against $\beta = 0$ is

$$\varphi_2(Z_{n-k+1:n}) := 1_{(-\infty,-u_\alpha)}(k^{1/2}(Z_{n-k+1:n} - \log(n/k))).$$

The asymptotic power functions of these asymptotic optimal tests are also provided by Theorem 2.1. By LeCam's first and third lemma we obtain that $\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=1}^k)$ and $\mathcal{L}_0((Z_{n-i+1:n})_{i=1}^k)$ are *contiguous* distributions and that under $\beta_n = \beta_n(\vartheta)$

$$\begin{aligned}\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_0((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1 \\ = \vartheta k^{1/2}(Z_{n-k+1:n} - \log(n/k)) + \vartheta^2/2 + o_{P_{\beta_n}}(1) \underset{\mathcal{D}_{\beta_n}}{\longrightarrow} N(\vartheta^2/2, \vartheta^2).\end{aligned}$$

As a consequence we obtain for the asymptotic power functions of φ_i

$$\lim_{n\to\infty} E_{\beta_n}(\varphi_i(Z_{n-k+1:n})) = 1 - \Phi(u_\alpha - |\vartheta|), \qquad i = 1, 2.$$

PROOF OF THEOREM 2.1. First observe that under $\beta = 0$ with $\vartheta \neq 0$

$$\begin{aligned}&P_0\{Z_{n:n} < 1/|\beta_n|^{1/2}\} \\ &\quad = P_0\{Z_{n:n} - \log(n) < k^{1/4}\log(n/k)/(2|\vartheta|)^{1/2} - \log(n)\} \to_{n\to\infty} 1\end{aligned}$$

since $P_0\{Z_{n:n} - \log(n) < x\} \to \exp(-e^{-x})$, $x \in \mathbb{R}$, and $k^{1/4}\log(n/k)/(2|\vartheta|)^{1/2} - \log(n) \to \infty$ as $n \to \infty$. As a consequence, we may suppose in the following that $Z_{n:n} < 1/|\beta_n|^{1/2}$. From formula (1.4.8) in Reiss (1989) we obtain in this case

$$\begin{aligned}(2.1)\quad &\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_0((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1 \\ &= \sum_{i=1}^k \log\left\{\frac{f_{\beta_n}(Z_{n-i+1:n})}{f_0(Z_{n-i+1:n})}\right\} + (n-k)\log\left\{\frac{F_{\beta_n}(Z_{n-k+1:n})}{F_0(Z_{n-k+1:n})}\right\} \\ &= \sum_{i=1}^k \left(Z_{n-i+1:n} - \frac{1+\beta_n}{\beta_n}\log\{1+\beta_n Z_{n-i+1:n}\}\right) \\ &\quad + (n-k)\log\left\{\frac{H_{\beta_n}(Z_{n-k+1:n})}{H_0(Z_{n-k+1:n})}\right\},\end{aligned}$$

provided $Z_{n-k+1:n} \geq x_0(\beta_n)$. But since $X_{(k)} := k^{1/2}(Z_{n-k+1:n} - \log(n/k))$ converges in distribution under $\beta = 0$ to the standard normal distribution and $x_0(\beta_n) - \log(n/k) < -\varepsilon < 0$ for some $\varepsilon > 0$ and all large n, the probability $P_0\{Z_{n-k+1:n} \geq x_0(\beta_n)\}$ of this event converges to one. In the following we suppress the index n of β_n. Recall that we suppose implicitly $Z_{n:n} < 1/|\beta_n|^{1/2}$. The following formula follows from the expansions $\log(1+\varepsilon) = \varepsilon - \varepsilon^2/2 + \varepsilon^3/3 + O(\varepsilon^4)$ and $\exp(\varepsilon) = 1 + \varepsilon + \varepsilon^2/2 + O(\varepsilon^3)$ for $\varepsilon \to 0$.

$$(2.2)\qquad H_\beta(Z_{n-k+1:n})/H_0(Z_{n-k+1:n}) - 1 = O_{P_0}(k^{1/2}/n).$$

From (2.2) we obtain that

$$\begin{aligned}
&(n-k)\log\{H_\beta(Z_{n-k+1:n})/H_0(Z_{n-k+1:n})\}\\
&\quad=\frac{(n-k)k\exp\{-X_{(k)}/k^{1/2}\}}{n-k\exp\{-X_{(k)}/k^{1/2}\}}\\
&\qquad\cdot\left(-\frac{\beta Z^2_{n-k+1:n}}{2}-\frac{\beta^2 Z^4_{n-k+1:n}}{8}+\frac{\beta^2 Z^3_{n-k+1:n}}{3}+o_{P_0}(k^{-3/2})\right)\\
&\qquad+O_{P_0}(k/n).
\end{aligned}$$

Observe now that $k\beta^2 Z^4_{n-k+1:n}/8$ converges in P_0-probability to $\vartheta^2/2$, as $Z_{n-k+1:n}/\log(n/k)$ converges to one. The right-hand side of the preceding equality equals therefore

$$\frac{(n-k)k\exp\{-X_{(k)}/k^{1/2}\}}{n-k\exp\{-X_{(k)}/k^{1/2}\}}\left(-\frac{\beta Z^2_{n-k+1:n}}{2}+\frac{\beta^2 Z^3_{n-k+1:n}}{3}\right)-\frac{\vartheta^2}{2}+o_{P_0}(1).$$

Consequently, we obtain

$$\begin{aligned}
(2.3)\qquad &(n-k)\log\left\{\frac{H_\beta(Z_{n-k+1:n})}{H_0(Z_{n-k+1:n})}\right\}+k\frac{\beta}{2}Z^2_{n-k+1:n}-k\frac{\beta^2}{3}Z^3_{n-k+1:n}\\
&=\left(k\frac{\beta}{2}Z^2_{n-k+1:n}-k\frac{\beta^2}{3}Z^3_{n-k+1:n}\right)\frac{1-\exp\{-X_{(k)}/k^{1/2}\}}{1-\dfrac{k}{n}\exp\{-X_{(k)}/k^{1/2}\}}\\
&\qquad-\frac{\vartheta^2}{2}+o_{P_0}(1)\\
&=\vartheta X_{(k)}-\vartheta^2/2+o_{P_0}(1).
\end{aligned}$$

In order to prove Theorem 2.1, it remains by (2.1) and (2.3) to show that

$$\begin{aligned}
(2.4)\qquad &\sum_{i=1}^k\Bigg(Z_{n-i+1:n}-\frac{1+\beta}{\beta}\log\{1+\beta Z_{n-i+1:n}\}\\
&\qquad-\frac{\beta}{2}Z^2_{n-k+1:n}+\frac{\beta^2}{3}Z^3_{n-k+1:n}\Bigg)=o_{P_0}(1).
\end{aligned}$$

The expansion $\log(1+\varepsilon)=\varepsilon-\varepsilon^2/2+\varepsilon^3/3+O(\varepsilon^4)$ implies that the left-hand side of (2.4) equals

$$\begin{aligned}
&\sum_{i=1}^k\Bigg(Z_{n-i+1:n}\\
&\qquad-(1+\beta)\left\{Z_{n-i+1:n}-\frac{\beta}{2}Z^2_{n-i+1:n}+\frac{\beta^2}{3}Z^3_{n-i+1:n}+O(\beta^3 Z^4_{n-i+1:n})\right\}\\
&\qquad-\frac{\beta}{2}Z^2_{n-k+1:n}+\frac{\beta^2}{3}Z^3_{n-k+1:n}\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{k-1} \Bigg(-\beta Z_{n-i+1:n} \\
&\qquad + (1+\beta)\left\{ \frac{\beta}{2} Z_{n-i+1:n}^2 - \frac{\beta^2}{3} Z_{n-i+1:n}^3 + O(\beta^3 Z_{n-i+1:n}^4) \right\} \\
&\qquad - \frac{\beta}{2} Z_{n-k+1:n}^2 + \frac{\beta^2}{3} Z_{n-k+1:n}^3 \Bigg) + o_{P_0}(1).
\end{aligned}
$$

Under the exponential distribution, the conditional distribution of $(Z_{n-i+1:n})_{i=1}^{k-1}$ given $Z_{n-k+1:n} = u > 0$, equals the distribution of $(W_{i:k-1} + u)_{i=k-1}^{1}$, where $W_1, W_2, \dots, W_{k-1}$ are independent and standard exponential rvs. This follows from Theorem 1.8.1 in Reiss (1989) and elementary computations. The conditional distribution of the right-hand side of the preceding equation given $Z_{n-k+1:n} = u \in I_n := [\log(n/k) - \varepsilon, \log(n/k) + \varepsilon]$, equals therefore the distribution of

$$
\begin{aligned}
&\sum_{i=1}^{k-1} \left(-\beta(W_i + u) + \frac{\beta}{2}((W_i + u)^2 - u^2) + \frac{\beta^2}{3}(u^3 - (W_i + u)^3) \right) \\
&\qquad + \beta \sum_{i=1}^{k-1} \left(\frac{\beta}{2}(W_i + u)^2 - \frac{\beta^2}{3}(W_i + u)^3 \right) + O\left(\beta^3 \sum_{i=1}^{k-1} (W_i + u)^4 \right) \\
&=: \sum_{i=1}^{k-1} \left(-\beta(W_i + u) + \frac{\beta}{2}((W_i + u)^2 - u^2) \right) + R_n(u), \\
&= \beta \sum_{i=1}^{k-1} \left((1 - W_i)(1 - u) + \left(\frac{W_i^2}{2} - 1 \right) \right) + R_n(u),
\end{aligned}
$$

where it is easy to see that $\sup_{u \in I_n} P_0\{|R_n(u)| > \varepsilon\} \to 0$ as $n \to \infty$, for any $\varepsilon > 0$ and by Tschebyscheff's inequality that also

$$
\sup_{u \in I_n} P_0 \left\{ \left| \sum_{i=1}^{k-1} \left(-\beta(W_i + u) + \frac{\beta}{2}((W_i + u)^2 - u^2) \right) \right| > \varepsilon \right\} \to_{u \to \infty} 0.
$$

This completes the proof of (2.4) and thus, also the proof of Theorem 2.1. □

The preceding result remains true, if we replace the condition that the upper tail of F *coincides* with the upper tail of a GPD by the assumption that it is in a *δ-neighborhood* of a GPD. To be precise, if we require that $\omega(F_\beta) = \omega(H_\beta)$ and

(2.5) $$|f_\beta(x)/h_\beta(x) - 1| \le C(1 - H_\beta(x))^\delta, \qquad x \ge x_0(\beta)$$

for some fixed $\delta, C > 0$ and $\limsup_{n\to\infty}(x_0(\beta_n) - \log(n/k)) < 0$, then we have again under $\beta = 0$ the expansion

$$
\begin{aligned}
&\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^{1}) / d\mathcal{L}_0((Z_{n-i+1:n})_{i=k}^{1})\}(Z_{n-i+1:n})_{i=k}^{1} \\
&\qquad = \vartheta k^{1/2}(Z_{n-k+1:n} - \log(n/k)) - \vartheta^2/2 + o_{P_0}(1) \xrightarrow[\mathcal{D}_0]{} N(-\vartheta^2/2, \vartheta^2),
\end{aligned}
$$

provided the sequences $k = k(n)$ satisfies the additional assumption $k(k/n)^\delta \to 0$ as $n \to \infty$. This can be shown along the lines of the preceding proof; for the sake of a clear presentation we omit the details.

The preceding result reveals that the complete information about the underlying parameter, which is contained in the vector of the upper k order statistics, is asymptotically already contained in the single order statistic $Z_{n-k+1:n}$. In order to know the consequences of this result for small up to moderate sample sizes n, we have to know about the rate at which $Z_{n-k+1:n}$ becomes asymptotically sufficient. To this end, we compare the distribution of $(Z_{n-i+1:n})_{i=k}^1$ with the distribution of $(Z_{n-k+1:n}, W_{(1)}, \dots, W_{(k-1)})$, where $W_{(1)}, \dots, W_{(k-1)}$ are generated in a two-step procedure. Given $Z_{n-k+1:n} = u$, generate $k-1$ independent standard exponential rvs $W_1, \dots, W_{k-1}$ and put

$$W_{(i)}(u) := u + W_{i:k-1}, \qquad 1 \le i \le k-1.$$

The rvs $W_{(i)}$ are then defined by

$$W_{(i)} := W_{(i)}(Z_{n-k+1:n}) = Z_{n-k+1:n} + W_{i:k-1}, \qquad 1 \le i \le k-1.$$

The motivation for the definition of this two step procedure is the fact that, if Z follows exactly a standard exponential distribution, the conditional distribution of $(Z_{n-i+1:n})_{i=k-1}^1$ given $Z_{n-k+1:n} = u > 0$ *equals* the distribution of $(u+W_{i:k-1})_{i=1}^{k-1}$ (see Theorem 1.8.1 in Reiss (1989)).

Clearly, $(Z_{n-k+1:n}, W_{(1)}, \dots, W_{(k-1)})$ carries only that information about the underlying parameter β which is contained in $Z_{n-k+1:n}$. The distance between the distributions of $(Z_{n-i+1:n})_{i=k}^1$ and of $(Z_{n-k+1:n}, W_{(1)}, \dots, W_{(k-1)})$ is then an upper bound for the lack of information in $Z_{n-k+1:n}$, compared with the complete information contained in $(Z_{n-i+1:n})_{i=k}^1$. Notice that our definition of asymptotic sufficiency is in the sense of that given in Proposition 2, Section 5.3 in the book by LeCam and Yang (1990).

Conditionally on $Z_{n-k+1:n}$, we deal with iid rvs and therefore, the *Hellinger distance* is the adequate distance between probability distributions to be considered here (see Section 3.3 of Reiss (1989) for details). Precisely, let Q_1, Q_2 be probability measures on the same measurable space and let μ be any measure dominating Q_1 and Q_2. The Hellinger distance $H(Q_1, Q_2)$ between Q_1 and Q_2 is then defined by

$$H(Q_1, Q_2) := \left(\int (f_1^{1/2} - f_2^{1/2})^2 d\mu \right)^{1/2},$$

where f_i is a μ density of Q_i, $i = 1, 2$. Note that the variational distance is bounded by the Hellinger distance.

THEOREM 2.2. *Choose $k = k(n) \in \{1, \dots, n\}$ such that $k \to \infty$, $k/n \to 0$ as $n \to \infty$. Put for $\vartheta \in \mathbb{R}$*

$$\beta_n := \beta_n(\vartheta) := \vartheta k^{-1/2} / \log^2(n/k).$$

Suppose that $\omega(F_{\beta_n}) = \omega(H_{\beta_n})$ *and*

$$|f_{\beta_n}(y)/h_{\beta_n}(y) - 1| \le C(1 - H_{\beta_n}(y))^\delta, \qquad y \ge x_0(\beta_n) \tag{2.6}$$

for some $C, \delta > 0$, *where* $x_0(\beta_n) - \log(n/k) \to_{n\to\infty} -\infty$. *Then*

$$\begin{aligned} H(\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1), \mathcal{L}_{\beta_n}(Z_{n-k+1:n}, (W_{(i)})_{i=1}^{k-1})) \\ = O(1/\log(n/k) + k^{1/2}(k/n)^\delta + \exp(-k^{1/2})). \end{aligned}$$

Remarks. One can show that the bound in the preceding result is sharp. Theorem 2.2 indicates therefore that the rate at which $Z_{n-k+1:n}$ becomes asymptotically sufficient is quite poor. The advice, to drop all the information contained in $(Z_{n-i+1:n})_{i=1}^{k-1}$ and to work with $Z_{n-k+1:n}$ alone, which is suggested by Theorem 2.1, can therefore be taken only with a grain of salt.

PROOF. By Theorem 1.8.1 in Reiss (1989), the distribution of $(Z_{n-i+1:n})_{i=k}^1$ can be generated by a two step procedure. First, generate $Z_{n-k+1:n} = z$ and then generate iid rvs $Y_1^{(z)}, \dots, Y_{k-1}^{(z)}$ with common df

$$F^{(z)}(t) := (F(t) - F(z))/(1 - F(z)), \qquad t \ge z.$$

The distribution of the vector $(Z_{n-k+1:n}, Y_{1:k-1}^{(\cdot)}, \dots, Y_{k-1:k-1}^{(\cdot)}) = T(Z_{n-k+1:n}, Y_1^{(\cdot)}, \dots, Y_{k-1}^{(\cdot)})$ then coincides with the distribution of $(Z_{n-i+1:n})_{i=k}^1$. By $T : \mathbb{R}^k \to \mathbb{R}^k$ we denote that functional which maps a vector $(x_1, \dots, x_k) \in \mathbb{R}^k$ onto the vector of its ordered values $(x_{1:k}, \dots, x_{k:k})$. Equally, we can write $(Z_{n-k+1:n}, W_{(1)}, \dots, W_{(k-1)}) = T(Z_{n-k+1:n}, Z_{n-k+1:n} + W_1, \dots, Z_{n-k+1:n} + W_{k-1})$, where $W_1, \dots, W_{k-1}$ are independent standard exponential rvs which are also independent of $Z_{n-k+1:n}$.

By the monotonicity theorem and the convexity theorem for the Hellinger distance (cf. Corollary 1.4.2 and Lemma 3.1.3 in Reiss (1993)) we obtain

$$\begin{aligned} &H^2(\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1), \mathcal{L}_{\beta_n}(Z_{n-k+1:n}, (W_{(i)})_{i=1}^{k-1})) \\ &\quad \le \int H^2(\mathcal{L}_{\beta_n}(z, Y_1^{(z)}, \dots, Y_{k-1}^{(z)}), \mathcal{L}_{\beta_n}(z, z + W_1, \dots, z + W_{k-1})) \\ &\qquad\qquad \mathcal{L}_{\beta_n}(Z_{n-k+1:n})(dz). \end{aligned} \tag{2.7}$$

For the sake of a clear presentation, we will drop in the following the index n of β_n.

By using the probability integral transformation and writing $Z_{n-k+1:n} = F_\beta^{-1}(U_{n-k+1:n})$, where $U_{n-k+1:n}$ is the kth-largest order statistic in an iid sample of n uniformly on $(0,1)$ distributed rvs, the following formula follows from the exponential inequality given in Lemma 3.1.1 in Reiss (1989) and elementary computations.

$$P_\beta\{|Z_{n-k+1:n} - \log(n/k)| > K\} = O(\exp(-k^{1/2})) \tag{2.8}$$

for $K > 0$ large enough.

From (2.7) and (2.8) we obtain for large n

$$
\begin{aligned}
(2.9)\quad & H^2(\mathcal{L}_\beta((Z_{n-i+1:n})_{i=k}^1), \mathcal{L}_\beta(Z_{n-k+1:n}, (W_{(i)})_{i=1}^{k-1})) \\
& \le \int_{\log(n/k)-K}^{\log(n/k)+K} H^2(\mathcal{L}_\beta(z, Y_1^{(z)}, \ldots, Y_{k-1}^{(z)}), \\
& \qquad\qquad \mathcal{L}_\beta(z, z+W_1, \ldots, z+W_{k-1}))\mathcal{L}_\beta(Z_{n-k+1:n})(dz) \\
& \quad + O(\exp(-k^{1/2})) \\
& \le k \int_{\log(n/k)-K}^{\log(n/k)+K} H^2(\mathcal{L}_\beta(Y_1^{(z)}), \mathcal{L}_\beta(z+W_1))\mathcal{L}_\beta(Z_{n-k+1:n})(dz) \\
& \quad + O(\exp(-k^{1/2})),
\end{aligned}
$$

where for $z \in [\log(n/k) - K, \log(n/k) + K] \subset (0, \infty)$ if n is large

$$
\begin{aligned}
& H^2(\mathcal{L}_\beta(Y_1^{(z)}), \mathcal{L}_\beta(z + W_1)) \\
& \quad = (1 - F_\beta(z))^{-1} \int_0^\infty (f_\beta^{1/2}(y+z)e^{y/2} - (1 - F_\beta(z))^{1/2})^2 e^{-y} dy.
\end{aligned}
$$

From condition (2.6) and elementary computations we obtain

$$
(2.10)\qquad |(1 - F_\beta(y))/(1 - H_\beta(y)) - 1| \le C_1(1 - H_\beta(y))^\delta
$$

for $y \in [x_0(\beta), \omega(H_\beta))$ and some constant $C_1 > 0$ not depending on β. And this implies

$$
(2.11)\qquad \int_{1/|\beta|^{1/2}}^\infty (f_\beta^{1/2}(y+z)e^{y/2} - (1 - F_\beta(z))^{1/2})^2 e^{-y} dy = O((k/n)^3)
$$

uniformly for $z \in [\log(n/k) - K, \log(n/k) + K]$.

If we show that

$$
\begin{aligned}
(2.12)\qquad & \frac{k}{1 - F_\beta(z)} \int_0^{1/|\beta|^{1/2}} (f_\beta^{1/2}(y+z)e^{y/2} - (1 - F_\beta(z))^{1/2})^2 e^{-y} dy \\
& \quad = O\left(\frac{1}{\log^2(n/k)} + k(k/n)^{2\delta}\right)
\end{aligned}
$$

uniformly for $z \in [\log(n/k) - K, \log(n/k) + K]$, the assertion of Theorem 2.2 then follows from (2.7), (2.9) and (2.11) by observing that by (2.10) $(1 - F_\beta(z))^{-1} = O(n/k)$ uniformly for $z \in [\log(n/k) - K, \log(n/k) + K]$. It remains therefore to prove (2.12).

By (2.6), (2.10) and the expansion $\log(1 + \varepsilon) = \varepsilon - \varepsilon^2/2 + O(\varepsilon^3)$ we obtain

$$
\int_0^{1/|\beta|^{1/2}} [f_\beta^{1/2}(y+z)e^{y/2} - (1 - F_\beta(z))^{1/2}]^2 e^{-y} dy
$$

$$= \int_0^{1/|\beta|^{1/2}} \Big[\exp\Big\{ \frac{y}{2} - \frac{1+\beta}{2\beta} \log(1+\beta(y+z)) \Big\} (1 + O((k/n)^\delta)) - \exp\Big\{ -\frac{1}{2\beta} \log(1+\beta z) \Big\} (1 + O((k/n)^\delta)) \Big]^2 e^{-y} dy$$
$$= O(\exp(-z)) \cdot \int_0^{1/|\beta|^{1/2}} [O\{|\beta|(y^2+2yz) + |\beta|(y+z) + \beta^2(y+z)^3 + (k/n)^\delta\}]^2 e^{-y} dy$$
$$= O(\exp(-z)(|\beta|^2 z^2 + (k/n)^{2\delta})) = O((k/n)(k^{-1}/\log^2(n/k) + (k/n)^{2\delta}))$$

which implies (2.12). □

Next we will consider the case $\beta_0 \neq 0$. For the sake of simplicity we drop in the following the von Mises parametrization of GPD's H_β for $\beta \neq 0$ and parametrize this subclass instead by

$$L_\beta(x) := \begin{cases} 1 - x^{-\beta}, & x \geq 1 \quad \text{if } \beta > 0 \\ 1 - (-x)^{-\beta}, & -1 \leq x \leq 0 \quad \text{if } \beta < 0. \end{cases}$$

Fix now $\beta_0 \neq 0$ and choose the alternatives $\beta_n = \beta_n(\vartheta)$ such that

$$(\beta_0 - \beta_n)/\beta_0 = \vartheta k^{-1/2}/\log(n/k), \qquad \vartheta \in \mathbb{R}.$$

For $F_\beta(x) = L_\beta(x)$, $x \geq x_0(\beta)$, with $\limsup_{n\to\infty} |x_0(\beta_n)|^{\beta_0} k/n < 1$, we have the following result.

THEOREM 2.3. (LAN) *For $\beta_0 \neq 0$ we have the expansion*

$$\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_{\beta_0}((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1$$
$$= \vartheta k^{1/2}(\beta_0 \log(|Z_{n-k+1:n}|) - \log(n/k))$$
$$- \vartheta^2/2 + o_{P_{\beta_0}}(1) \underset{\mathcal{D}_{\beta_0}}{\longrightarrow} N(-\vartheta^2/2, \vartheta^2).$$

Observe that if Z has df L_{β_0} with $\beta_0 \neq 0$, then $\beta_0 \log(|Z|)$ follows the standard exponential distribution.

The preceding result shows that also in case $\beta_0 \neq 0$, the complete information about the underlying parameter, which is contained in the vector of the kth-largest order statistics in the sample, is asymptotically already contained in the kth-largest order statistics $Z_{n-k+1:n}$ alone. The remarks after Theorem 2.1 on the derivation of optimal tests and the computation of their limiting power functions from the LAN expansion of the loglikelihood ratios carry over to Theorem 2.3.

This result gives also new insight into the problem of *estimating* the extreme value index β with $\beta > 0$. Consider the subclass $\{F_\beta : F_\beta(x) = L_{1/\beta}(x)$ for $x \geq$

$x_0(\beta), \beta > 0\}$ of dfs, whose upper tails ultimately coincide with the upper tails of a Pareto distribution. A popular estimator of β is the Hill (1975) estimator

$$\hat{\beta}_n(k) := (k-1)^{-1} \sum_{i=1}^{k-1} \log(Z_{n-i+1:n}) - \log(Z_{n-k+1:n})$$

which can easily be motivated by maximum likelihood theory. Observe that conditional on $Z_{n-k+1:n} = u$, $(\log(Z_{n-i+1:n}/Z_{n-k+1:n}))_{i=k-1}^{1}$ equals $(\beta W_{i:k-1})_{i=1}^{k-1}$ in distribution if u is large, where $W_1, W_2, \ldots$ are independent and standard exponential rvs. By this argument it is readily seen that

$$(k^{1/2}/\beta)(\hat{\beta}_n(k) - \beta) \xrightarrow[\mathcal{D}_\beta]{} N(0,1).$$

But the Hill estimator is outperformed by the estimator

$$\hat{b}_n := \log(|Z_{n-k+1:n}|)/\log(n/k)$$

as

$$(k^{1/2}\log(n/k)/\beta)(\hat{b}_n - \beta) \xrightarrow[\mathcal{D}_\beta]{} N(0,1).$$

This observation can be explained by the preceding result revealing the asymptotic sufficiency of $Z_{n-k+1:n}$. For the proof of asymptotic normality of Hill's (1975) estimator under general conditions we refer to Csörgő and Mason (1985), Hall and Welsh (1985), Smith (1987), and the literature cited therein.

Asymptotically optimal estimators of $\beta > 0$ under full and partial knowledge of the slowly varying function $\psi(x)$ at the tail in the model $F(x) = 1 - x^{-\beta}\psi(x)$, $x \geq x_0$, are proposed and investigated by Wei (1992). This paper provides also a useful survey of the literature on Hill's and related estimators of β.

Theorem 2.3 as well as the preceding remarks remain again true, if we replace the condition that the upper tail of F_β coincides with the upper tail of L_β by the condition that it is in a δ-neighborhood of L_β i.e., $\omega(F_\beta) = \omega(L_\beta)$ and

$$|f_\beta(x)/l_\beta(x) - 1| \leq C(1 - L_\beta(x))^\delta, \qquad x \geq x_0(\beta)$$

for some fixed $\delta, C > 0$, where l_β denotes the density of L_β. If $\limsup_{n\to\infty} |x_0(\beta_n)|^{\beta_0} k/n < 1$, then we have again under $\beta_0 \neq 0$ the expansion

$$\begin{aligned}
&\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^{1})/d\mathcal{L}_{\beta_0}((Z_{n-i+1:n})_{i=k}^{1})\}(Z_{n-i+1:n})_{i=k}^{1} \\
&\quad = \vartheta k^{1/2}(\beta_0 \log(|Z_{n-k+1:n}|) - \log(n/k)) \\
&\qquad - \vartheta^2/2 + o_{P_{\beta_0}}(1) \xrightarrow[\mathcal{D}_{\beta_0}]{} N(-\vartheta^2/2, \vartheta^2),
\end{aligned}$$

provided the sequence $k = k(n)$ satisfies the additional assumption $k(k/n)^\delta \to 0$ as $n \to \infty$. This can be shown along the lines of the preceding proof. In order not to overload the paper with too many technicalities we omit the details.

PROOF OF THEOREM 2.3. From formula (1.4.8) in Reiss (1989) we obtain

$$\begin{aligned}(2.13)\qquad &\log\{d\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_{\beta_0}((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1\\ &=\sum_{i=1}^k \log\{(\beta_n/\beta_0)|Z_{n-i+1:n}|^{\beta_0-\beta_n}\}\\ &\quad+(n-k)\log\{L_{\beta_n}(Z_{n-k+1:n})/L_{\beta_0}(Z_{n-k+1:n})\},\end{aligned}$$

provided $Z_{n-k+1:n}\ge x_0(\beta_n)$. But the probability of this event converges to one by the condition $\limsup_{n\to\infty}|x_0(\beta_n)|^{\beta_0}(k/n) < 1$ and the fact that $k^{1/2}(\beta_0\log(|Z_{n-k+1:n}|)-\log(n/k)) \xrightarrow[\mathcal{D}_0]{} N(0,1)$.

Recall that $\beta_0\log(|Z|)$ follows the standard exponential distribution if Z has df L_{β_0}, and put

$$Y_{n-i+1:n} := \beta_0\log(|Z_{n-i+1:n}|), \qquad 1\le i\le n.$$

Then the df of $Y_{n-i+1:n}$ coincides ultimately with the df of the ith-largest order statistics in a sample of n independent and standard exponential rvs. In particular we have therefore

$$X_{(k)} := k^{1/2}(Y_{n-k+1:n}-\log(n/k)) \xrightarrow[\mathcal{D}_{\beta_0}]{} N(0,1).$$

The next formula follows from the expansion $\exp(\varepsilon)-1=\varepsilon+\varepsilon^2/2+O(\varepsilon^3)$ as $\varepsilon\to 0$, the fact that $(\beta_0-\beta_n)/\beta_0 = \vartheta k^{-1/2}/\log(n/k)$ and elementary computations. Recall that we assume implicitly that $Z_{n-k+1:n} > x_0(\beta_n)$:

$$(2.14)\qquad L_{\beta_n}(Z_{n-k+1:n})/L_{\beta_0}(Z_{n-k+1:n})-1 = O_{P_{\beta_0}}(k^{1/2}/n).$$

From (2.14) by Taylor expansion of exp at zero and the fact that $Y_{n-k+1:n}/\log(n/k)\to 1$ in P_{β_0}-probability, we obtain the expansion

$$\begin{aligned}(2.15)\qquad &(n-k)\log\left\{\frac{L_{\beta_n}(Z_{n-k+1:n})}{L_{\beta_0}(Z_{n-k+1:n})}\right\}+\frac{\vartheta k^{1/2}}{\log(n/k)}Y_{n-k+1:n}\\ &=\vartheta X_{(k)}-\frac{\vartheta^2}{2}+o_{P_{\beta_0}}(1).\end{aligned}$$

The assertion of Theorem 2.3 follows now from (2.13) and (2.15), if we show that

$$(2.16)\qquad \sum_{i=1}^k \log\left\{\frac{\beta_n}{\beta_0}|Z_{n-i+1:n}|^{\beta_0-\beta_n}\right\}-\frac{\vartheta k^{1/2}}{\log(n/k)}Y_{n-k+1:n}=o_{P_{\beta_0}}(1).$$

But the left-hand side of (2.16) equals

$$\begin{aligned}&\sum_{i=1}^k\left\{\frac{\beta_0-\beta_n}{\beta_0}(Y_{n-i+1:n}-Y_{n-k+1:n})+\log\left(1+\frac{\beta_n-\beta_0}{\beta_0}\right)\right\}\\ &=\frac{\beta_0-\beta_n}{\beta_0}\sum_{i=1}^k\{Y_{n-i+1:n}-Y_{n-k+1:n}-1\}+o(1).\end{aligned}$$

Recall now that conditional on $Y_{n-k+1:n} = u$, the distribution of $(Y_{n-i+1:n})_{i=k-1}^{1}$ equals the distribution of $(W_{i:k-1} + u)_{i=1}^{k-1}$, where $W_1, W_2, \ldots, W_{k-1}$ are independent and standard exponential random variables. Consequently, conditional on $Y_{n-k+1:n} = u$, the right-hand side of the preceding equation equals in distribution

$$\frac{\vartheta k^{-1/2}}{\log(n/k)} \sum_{i=1}^{k-1}(W_i - 1) + o(1).$$

This implies (2.16) by conditioning on $Y_{n-k+1:n} = u \in [\log(n/k) - \varepsilon, \log(n/k) + \varepsilon]$ for some small $\varepsilon > 0$. □

In the following we will establish a bound for the rate at which $Z_{n-k+1:n}$ becomes asymptotically sufficient. This will be done by proving a result which parallels Theorem 2.2.

Suppose that $Z_1, \ldots, Z_n$ are iid rvs with common df L_β, $\beta \neq 0$. Then the vector $(Z_{n-i+1:n})_{i=k-1}^{1}$ given $Z_{n-k+1:n} = z$, equals in distribution the vector $(W_{i:k-1})_{i=1}^{k-1}$, where $W_1, \ldots, W_{k-1}$ are iid rvs with common df

$$(L_\beta(t) - L_\beta(z))/(1 - L_\beta(z)) = L_\beta(t/|z|), \qquad z \leq t < \omega(L_\beta)$$

(cf. Theorem 1.8.1 in Reiss (1989)). Consequently, conditional on $Z_{n-k+1:n} = z$, the vector $(Z_{n-i+1:n})_{i=k-1}^{1}$ equals in distribution the vector $(|z|U_{i:k-1})_{i=1}^{k-1}$, where $U_1, \ldots, U_k$ are iid rvs with common df L_β.

We compare in the following the vector $(Z_{n-i+1:n})_{i=k}^{1}$ with the vector $(Z_{n-k+1:n}, V_{(1)}, \ldots, V_{(k-1)})$, where the $V_{(i)}$ are generated by the following two step procedure, which is motivated by the preceding considerations.

Given $Z_{n-k+1:n} = z$, we generate $k-1$ iid rvs $V_1, \ldots, V_{k-1}$ with common df L_{β_0}, independent of the underlying parameter β, and define

$$V_{(i)}(z) := |z|V_{i:k-1}, \qquad 1 \leq i \leq k-1.$$

The rvs $V_{(1)}, \ldots, V_{(k-1)}$ are then defined by

$$V_{(i)} := V_{(i)}(Z_{n-k+1:n}) = |Z_{n-k+1:n}|V_{i:k-1}, \qquad 1 \leq i \leq k-1.$$

The vector $(Z_{n-k+1:n}, V_{(1)}, \ldots, V_{(k-1)})$ contains therefore only that information about the underlying parameter β which is contained in $Z_{n-k+1:n}$, and the Hellinger distance

$$H(\mathcal{L}_\beta((Z_{n-i+1:n})_{i=k}^{1}), \mathcal{L}_\beta(Z_{n-k+1:n}, V_{(1)}, \ldots, V_{(k-1)}))$$

is an upper bound for the lack of information in $Z_{n-k+1:n}$. Notice that this Hellinger distance is zero if $\beta = \beta_0$.

THEOREM 2.4. *Choose $k = k(n) \in \{1, \ldots, n\}$ such that $k \to \infty$, $k/n \to 0$ as $n \to \infty$. Fix $\beta_0 \neq 0$ and define for $\vartheta \in \mathbb{R}$ the sequence $\beta_n = \beta_n(\vartheta)$ by*

$$(\beta_0 - \beta_n)/\beta_0 = \vartheta k^{-1/2}/\log(n/k).$$

Suppose that $\omega(F_\beta)=\omega(L_\beta)$ *and that* F_β *ultimately has a density* f_β *such that*

$$|f_\beta(y)/l_\beta(y)-1|\le C(1-L_\beta(y))^\delta, \qquad y\in[x_0(\beta),\omega(F_\beta)), \tag{2.17}$$

for some fixed $C,\delta>0$, *with* $\lim_{n\to\infty}|x_0(\beta_n)|^{\beta_0}k/n=0$. *Then,*

$$\begin{aligned}H(\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1),\ &\mathcal{L}_{\beta_n}(Z_{n-k+1:n},V_{(1)},\dots,V_{(k)}))\\ &=O(1/\log(n/k)+k^{1/2}(k/n)^\delta+\exp(-k^{1/2})).\end{aligned}$$

PROOF. Repeating the arguments in the proof of Theorem 2.3 we obtain

$$\begin{aligned}H^2(\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1),\ &\mathcal{L}_{\beta_n}(Z_{n-k+1:n},V_{(1)},\dots,V_{(k-1)}))\\ &\le\int H^2(\mathcal{L}_{\beta_n}(z,X_1^{(z)},\dots,X_{k-1}^{(z)}),\mathcal{L}_{\beta_n}(z,|z|V_1,\dots,|z|V_{k-1}))\\ &\qquad\qquad\cdot\mathcal{L}_{\beta_n}(Z_{n-k+1:n})(dz),\end{aligned}$$

where $X_1^{(z)},\dots,X_{k-1}^{(z)}$ are iid with common df

$$F_{\beta_n}^{(z)}(t):=(F_{\beta_n}(t)-F_{\beta_n}(z))/(1-F_{\beta_n}(z)), \qquad t\ge z.$$

If z is large i.e., if $z\ge x_0(\beta_n)$, the df $F_{\beta_n}^{(z)}$ has density $f_{\beta_n}^{(z)}(t)=f_{\beta_n}(t)/(1-F_{\beta_n}(z))$, $t\ge z$. In complete analogy to the proof of (2.8) it is shown that for $\varepsilon>0$ small enough

$$P_{\beta_n}\{\varepsilon<|Z_{n-k+1:n}|^{\beta_n}k/n<1/\varepsilon\}=1+O(\exp(-k^{1/2})). \tag{2.18}$$

As the Hellinger distance is in general bounded by $\sqrt{2}$, we obtain from (2.18) and Lemma 3.3.10(i) of Reiss (1989)

$$\begin{aligned}H^2(\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1),\ &\mathcal{L}_{\beta_n}(Z_{n-k+1:n},V_{(1)},\dots,V_{(k-1)}))\\ &\le k\int_{\varepsilon<|z|^{\beta_n}k/n<1/\varepsilon}H^2(\mathcal{L}_{\beta_n}(X_1^{(z)}),\mathcal{L}_{\beta_n}(|z|V_1))\mathcal{L}_{\beta_n}(Z_{n-k+1:n})(dz)\\ &\quad+O(\exp(-k^{1/2})).\end{aligned}$$

Now for z such that $\varepsilon<|z|^{\beta_n}k/n<1/\varepsilon$ we have $x_0(\beta_n)<z<\omega(F_{\beta_n})$ if n is large; this follows from the condition $\lim_{n\to\infty}|x_0(\beta_n)|^{\beta_0}k/n=0$. Thus we can write for such z

$$\begin{aligned}&H^2(\mathcal{L}_{\beta_n}(X_1^{(z)}),\mathcal{L}_{\beta_n}(|z|V_1))\\ &\quad=(1-F_{\beta_n}(z))^{-1}\int_{\mathbb{R}}[(|z|f_{\beta_n}(t|z|)/l_{\beta_0}(t))^{1/2}-(1-F_{\beta_n}(z))^{1/2}]^2l_{\beta_0}(t)dt.\end{aligned}$$

We obtain therefore from (2.17) that uniformly for z with $\varepsilon < |z|^{\beta_n} k/n < 1/\varepsilon$

$$(2.19)\quad \int_{\mathbb{R}} \left[|z|^{1/2} \frac{f_{\beta_n}^{1/2}(t|z|)}{l_{\beta_0}^{1/2}(t)} - (1 - F_{\beta_n}(z))^{1/2} \right]^2 l_{\beta_0}(t)dt$$

$$= \int_{\mathbb{R}} \Bigg[|z|^{1/2} \frac{l_{\beta_n}^{1/2}(t|z|)}{l_{\beta_0}^{1/2}(t)} \left\{ 1 + \left(\frac{f_{\beta_n}^{1/2}(t|z|)}{l_{\beta_n}^{1/2}(t|z|)} - 1 \right) \right\} - (1 - L_{\beta_n}(z))^{1/2} \cdot \left\{ 1 + \left(\left(\frac{1 - F_{\beta_n}(z)}{1 - L_{\beta_n}(z)} \right)^{1/2} - 1 \right) \right\} \Bigg]^2 l_{\beta_0}(t)dt$$

$$= \int_{\mathbb{R}} \Bigg[|z|^{1/2} \frac{l_{\beta_n}^{1/2}(t|z|)}{l_{\beta_0}^{1/2}(t)} - (1 - L_{\beta_n}(z))^{1/2} + \left\{ |z|^{1/2} \frac{l_{\beta_n}^{1/2}(t|z|)}{l_{\beta_0}^{1/2}(t)} + (1 - L_{\beta_n}(z))^{1/2} \right\} \cdot O((1 - L_{\beta_n}(z))^{\delta}) \Bigg]^2 l_{\beta_0}(t)dt$$

$$= (1 - L_{\beta_n}(z)) \cdot \int_{\mathbb{R}} \Bigg[\left(\frac{\beta_n}{\beta_0} \right)^{1/2} |t|^{(\beta_0 - \beta_n)/2} - 1 + \left\{ \left(\frac{\beta_n}{\beta_0} \right)^{1/2} |t|^{(\beta_0 - \beta_n)/2} + 1 \right\} O((k/n)^{\delta}) \Bigg]^2 l_{\beta_0}(t)dt$$

as $|z|^{-\beta_n/2} = 1 - L_{\beta_n}(z)$. The preceding integral equals

$$(2.20)\quad \int_{\mathbb{R}} [(1 + O(\beta_n - \beta_0))|t|^{(\beta_0 - \beta_n)/2} - 1 + O((k/n)^{\delta}(1 + |t|^{(\beta_0 - \beta_n)/2}))]^2 l_{\beta_0}(t)dt$$

$$= \int_{\mathbb{R}} [|t|^{(\beta_0 - \beta_n)/2} - 1 + O((k/n)^{\delta}) + O(\{|\beta_n - \beta_0| + (k/n)^{\delta}\}|t|^{(\beta_0 - \beta_n)/2})]^2 l_{\beta_0}(t)dt.$$

Observe now that

$$\int_{\mathbb{R}} [|t|^{(\beta_0 - \beta_n)} - 1]^2 l_{\beta_0}(t)dt = \frac{(\beta_0 - \beta_n)^2}{\beta_n(\beta_0 - \beta_n)} = O(k^{-1}/\log^2(n/k))$$

and that

$$\int_{\mathbb{R}} |t|^{\beta_0 - \beta_n} l_{\beta_0}(t)dt = O(1)$$

if n is large, since $\beta_0 - \beta_n \to 0$ as $n \to \infty$. Consequently, we obtain

$$\text{(2.21)} \qquad \int_{\mathbb{R}} [|t|^{(\beta_0-\beta_n)/2} - 1 + O((k/n)^\delta) + O(\{|\beta_n - \beta_0| + (k/n)^\delta\}|t|^{(\beta_0-\beta_n)/2})]^2 l_{\beta_0}(t)dt = O(k^{-1}/\log^2(n/k) + (k/n)^{2\delta}).$$

From (2.19)-(2.21) we get

$$\begin{aligned} &H^2(\mathcal{L}_{\beta_n}((Z_{n-i+1:n})_{i=k}^1), \mathcal{L}_{\beta_n}(Z_{n-k+1:n}, V_{(1)}, \dots, V_{k-1})) \\ &\le k \int_{\varepsilon<|z|^{\beta_n}(k/n)<\varepsilon^{-1}} H^2(\mathcal{L}_{\beta_n}(X_n^{(z)}), \mathcal{L}_{\beta_n}(|z|V_1))\mathcal{L}_{\beta_n}(Z_{n-k+1:n})(dz) \\ &\quad + O(\exp(-k^{1/2})) \\ &\le k \int_{\varepsilon<|z|^{\beta_n}(k/n)<\varepsilon^{-1}} \frac{1 - L_{\beta_n}(z)}{1 - F_{\beta_n}(z)} O\left(\frac{k^{-1}}{\log^2(n/k)} + \left(\frac{k}{n}\right)^{2\delta}\right) L_{\beta_n}(Z_{n-k+1:n})(dz) \\ &\quad + O(\exp(-k^{1/2})) \\ &= O(\log^{-2}(n/k) + k(k/n)^{2\delta} + \exp(-k^{1/2})). \end{aligned}$$

This completes the proof of Theorem 2.4. □

3. Adding a scale and location parameter

In the following we will extend the model (M) and require

$$F(x) = F_\beta(cx + d) =: F_{\beta,c,d}(x) \qquad \text{for all} \quad x \ge x_0 = x_0(\beta, c, d),$$

for some $c > 0$, $d \in \mathbb{R}$ with $\{F_\beta : \beta \in \Theta\}$ being a parametric family of dfs. We suppose again that $x_0 = x_0(\beta, c, d)$ is unknown. Our testing problem is now

$$\mathcal{L}_{\beta_0,c_0,d_0}((Z_{n-i+1:n})_{i=k}^1) \quad \text{against} \quad \mathcal{L}_{\beta_n,c_n,d_n}((Z_{n-i+1:n})_{i=k}^1)$$

where $Z_1, \dots, Z_n$ are iid rvs with common df $F_{\beta,c,d}$.

Without loss of generality we assume that $c_0 = 1$ and $d_0 = 0$, as this can be achieved by the data transformation $Z_i' := c_0 Z_i + d_0$, $1 \le i \le n$. We consider again the particular parametric family

$$F_\beta(x) = H_\beta(x) = 1 - (1 + \beta x)^{-1/\beta}, \qquad 0 < (1 + \beta x)^{-1/\beta} < 1,$$

of GPDs in their von Mises parametrization, and we will test at first the exponential distribution $\beta = 0$. Precisely, put for $\vartheta, \xi, \eta \in \mathbb{R}$

$$\begin{aligned} \beta_n &:= \beta_n(\vartheta) := 2\vartheta k^{-1/2}/\log^2(n/k), \\ c_n &:= c_n(\xi) := 1 - \xi k^{-1/2}/\log(n/k), \\ d_n &:= d_n(\eta) := -\eta k^{-1/2}, \end{aligned}$$

where the sequence $k = k(n) \in \{1, \dots, n\}$ satisfies $k \to \infty$ but $k/n \to 0$ as $n \to \infty$.

THEOREM 3.1. (LAN) *Suppose that* $\limsup_{n\to\infty}(x_0(\beta_n, c_n, d_n) - \log(n/k)) < 0$. *Then we have under the hypothesis* $(\beta, c, d) = (0, 1, 0)$ *for any* $(\vartheta, \xi, \eta) \in \mathbb{R}^3$

$$\begin{aligned}
&\log\{d\mathcal{L}_{\beta_n,c_n,d_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_{0,1,0}((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1 \\
&\quad= (\vartheta + \xi + \eta)k^{1/2}(Z_{n-k+1:n} - \log(n/k)) - (\vartheta + \xi + \eta)^2/2 + o_{P_{0,1,0}}(1) \\
&\quad\underset{\mathcal{D}_{0,1,0}}{\longrightarrow} N(-(\vartheta + \xi + \eta)^2/2, (\vartheta + \xi + \eta)^2).
\end{aligned}$$

The preceding result reveals that the kth-largest order statistic $Z_{n-k+1:n}$ remains the central sequence for the loglikelihood processes pertaining to the vector $(Z_{n-i+1:n})_{i=1}^k$, if we add an unknown scale and location parameter. As in the discussion after Theorem 2.1, asymptotically optimal tests for (β, c, d) can therefore be based on $Z_{n-k+1:n}$. But if $\vartheta + \xi + \eta = 0$, that is, if the vector $(\vartheta, \xi, \eta) \in \mathbb{R}^3$ is on the hyperplane generated by the basis $(1, -1, 0)$, $(0, -1, 1)$, then the preceding result shows that one cannot distinguish asymptotically between $(\beta_n(\vartheta), c_n(\xi), d_n(\eta))$ and $(0, 1, 0)$. This means that an alternative shape parameter $\beta_n(\vartheta)$ can be hidden by a scale and location parameter, such that hypothesis and alternative cannot be separated asymptotically.

Theorem 3.1 remains again true, if we replace the condition that the upper tail of F_β coincides with that of a GPD by the condition that it is in a δ-neighborhood of a GPD: If we require $\omega(F_\beta) = \omega(H_\beta)$ and that F_β has ultimately a density f_β with

$$|f_\beta(x)/h_\beta(x) - 1| \le C(1 - H_\beta(x))^\delta, \qquad x \ge x_0(\beta) \tag{3.1}$$

for some fixed $\delta, C > 0$ and $\limsup_{n\to\infty}(x_0(\beta_n) - \log(n/k)) < 0$, then we have under $\beta = 0$, $c = 1$, $d = 0$ again the expansion

$$\begin{aligned}
&\log\{d\mathcal{L}_{\beta_n,c_n,d_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_{0,1,0}((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1 \\
&\quad= (\vartheta + \xi + \eta)k^{1/2}(Z_{n-k+1:n} - \log(n/k)) - (\vartheta + \xi + \eta)^2/2 + o_{P_{0,1,0}}(1) \\
&\quad\underset{\mathcal{D}_{0,1,0}}{\longrightarrow} N(-(\vartheta + \xi + \eta)^2/2, (\vartheta + \xi + \eta)^2),
\end{aligned}$$

provided $k(k/n)^\delta \to 0$ as $n \to \infty$.

PROOF OF THEOREM 3.1. Put $X_{(k)} := k^{1/2}(Z_{n-k+1:n} - \log(n/k))$. In complete analogy to the proof of (2.2) one establishes

$$H_{\beta_n,c_n,d_n}(Z_{n-k+1:n})/H_0(Z_{n-k+1:n}) - 1 = O_P(k^{1/2}/n). \tag{3.2}$$

From now on we will drop the index n of β_n etc. As in the proof of Theorem 2.1 we have

$$\begin{aligned}
&\log\{d\mathcal{L}_{\beta,c,d}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_{0,1,0}((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1 \\
&\quad= \sum_{i=k}^{k}\left(Z_{n-i+1:n} - \frac{1+\beta}{\beta}\log(1 + \beta(cZ_{n-i+1:n} + d)) - \log(c)\right) \\
&\qquad+ (n-k)\log\{H_{\beta,c,d}(Z_{n-k+1:n})/H_0(Z_{n-k+1:n})\},
\end{aligned} \tag{3.3}$$

provided $Z_{n-k+1:n} > x_0(\beta, c, d)$. But the probability of this event converges to one as $\limsup_{n\to\infty}(x_0(\beta, c, d) - \log(n/k)) < 0$ and $k^{1/2}(Z_{n-k+1:n} - \log(n/k)) \underset{\mathcal{D}_{0,1,0}}{\longrightarrow} N(0,1)$.

By using (3.2) and the expansion $\log(1+\varepsilon) = \varepsilon + O(\varepsilon^2)$ it is elementary to show that

$$(3.4) \qquad \begin{aligned} &(n-k)\log\{H_{\beta,c,d}(Z_{n-k+1:n})/H_0(Z_{n-k+1:n})\} \\ &\quad - k\Big(d - (1-c)Z_{n-k+1:n} \\ &\qquad - \frac{\beta}{2}(cZ_{n-k+1:n} + d)^2 + \frac{\beta^2}{3}(cZ_{n-k+1:n} + d)^3\Big) \\ &= (\vartheta + \xi + \eta)X_{(k)} - (\vartheta + \xi + \eta)^2/2 + o_{P_{0,1,0}}(1). \end{aligned}$$

By using the expansion $\log(1+\varepsilon) = \varepsilon - \varepsilon^2/2 + \varepsilon^3/3 + O(\varepsilon^4)$ and the conditioning technique in the proof of formula (2.1) one shows

$$(3.5) \qquad \begin{aligned} &\sum_{i=1}^{k}\Big(Z_{n-i+1:n} - \frac{1+\beta}{\beta}\log(1 + \beta(cZ_{n-i+1:n} + d)) - \log(c)\Big) \\ &\quad + k\Big(d - (1-c)Z_{n-k+1:n} \\ &\qquad - \frac{\beta}{2}(cZ_{n-k+1:n} + d)^2 + \frac{\beta^2}{3}(cZ_{n-k+1:n} + d)^3\Big) \\ &= o_{P_{0,1,0}}(1). \end{aligned}$$

The assertion of Theorem 3.1 is then immediate from (3.3)-(3.5). □

Next we will consider the case $\beta_0 \neq 0$ with underlying tail distributions of the form

$$L_{\beta,c}(x) := L_\beta(cx) = \begin{cases} 1 - (cx)^{-\beta}, & cx \geq 1, & \text{if } \beta > 0 \\ 1 - (-cx)^{-\beta}, & -1 \leq cx \leq 0, & \text{if } \beta < 0 \end{cases}$$

and $c > 0$. Fix $\beta_0 \neq 0$ and choose with $\vartheta, \xi \in \mathbb{R}$ the alternatives $\beta_n = \beta_n(\vartheta)$, $c_n = c_n(\xi)$ of (β_0, c_0) with $c_0 = 1$ as

$$(\beta_0 - \beta_n)/\beta_0 = \vartheta k^{-1/2}/\log(n/k), \qquad 1 - c_n = \xi k^{-1/2}/\beta_0.$$

For $F_{\beta,c}(x) = L_{\beta,c}(x)$, $x \geq x_0(\beta, c)$, with $\limsup_{n\to\infty}|x_0(\beta_n, c_n)|^{\beta_0}k/n < 1$ we have the following result. Recall that $\beta_0\log(|Z|)$ has a standard exponential distribution if Z has df L_{β_0}. This implies the asymptotic normality of $\beta_0\log(|Z_{n-k+1:n}|)$ in the following result.

The following result parallels Theorem 3.1 in the case $\beta_0 \neq 0$. But note that a location parameter is missing, in which case $Z_{n-k+1:n}$ loses its asymptotic sufficiency; see Theorem 3.3 below.

THEOREM 3.2. (LAN) *For $\beta_0 \neq 0$ and $\beta_n = \beta_n(\vartheta)$, $c_n = c_n(\xi)$ as defined above we have*

$$\begin{aligned}&\log\{d\mathcal{L}_{\beta_n,c_n}((Z_{n-i+1:n})_{i=k}^{1})/d\mathcal{L}_{\beta_0,1}((Z_{n-i+1:n})_{i=k}^{1})\}(Z_{n-i+1:n})_{i=k}^{1}\\&\quad= (\vartheta+\xi)k^{1/2}(\beta_0\log(|Z_{n-k+1:n}|)-\log(n/k))-(\vartheta+\xi)^2/2+o_{P_{\beta_0,1}}(1)\\&\quad\xrightarrow[\mathcal{D}_{\beta_0,1}]{} N(-(\vartheta+\xi)^2/2,(\vartheta+\xi)^2).\end{aligned}$$

Theorem 3.2 remains true if we require that $\omega(F_\beta)=\omega(L_\beta)$, F_β has ultimately a density f_β such that

$$|f_\beta(x)/h_\beta(x)-1|\le C(1-L_\beta(x))^\delta,\qquad x\ge x_0(\beta) \tag{3.6}$$

for some fixed $\delta, C > 0$ with $\limsup_{n\to\infty}|x_0(\beta_n)|^{\beta_0}k/n < 1$, and $\lim_{n\to\infty}k(k/n)^\delta = 0$. Note that in case $\beta_0>0$ an additional location parameter can be subsumed under condition (3.6), but affecting the exponent δ. In case $\beta_0<0$, an additional location parameter alters the right endpoint of the pertaining dfs; see Theorem 3.3 below.

PROOF. Put $Y_{n-i+1:n} := \beta_0\log(|Z_{n-i+1:n}|)$, $1\le i\le n$ and $X_{(k)} := k^{1/2}(Y_{n-k+1:n}-\log(n/k)) \xrightarrow[\mathcal{D}_{\beta_0,1}]{} N(0,1)$. The assertion of Theorem 3.2 then follows in complete analogy to the arguments in the proof of Theorem 2.3 by the expansions

$$L_{\beta_n,c_n}(Z_{n-k+1:n})/L_{\beta_0,1}(Z_{n-k+1:n})-1=O_{P_{\beta_0,1}}(k^{1/2}/n)$$

and, by means of formula (2.16),

$$\begin{aligned}&\sum_{i=1}^{k}\log\left\{\frac{\beta_n}{\beta_0}|Z_{n-i+1:n}|^{\beta_0-\beta_n}c_n^{-\beta_n}\right\}-k\left(\frac{\beta_0-\beta_n}{\beta_0}Y_{n-k+1:n}-\beta_n\log(c_n)\right)\\&\quad=\sum_{i=1}^{k}\left\{\frac{\beta_0-\beta_n}{\beta_0}(Y_{n-i+1:n}-Y_{n-k+1:n})+\log\left(1+\frac{\beta_n-\beta_0}{\beta_0}\right)\right\}\\&\quad=\frac{\beta_0-\beta_n}{\beta_0}\sum_{i=1}^{k}\{Y_{n-i+1:n}-Y_{n-k+1:n}-1\}+o(1)=o_{P_{\beta_0,1}}(1).\end{aligned}$$ □

We complete this paper by showing that the kth-largest order statistic $Z_{n-k+1:n}$ is no longer the central sequence, if we add in the model $F_{\beta,c}(x) = L_\beta(cx)$, $x\ge x_0(\beta,c)$ with $|\beta|>2$ an unknown location parameter $d\in\mathbb{R}$, and consider

$$F_{\beta,c,d}(x)=L_{\beta,c,d}(x):=L_\beta(cx+d),\qquad x_0(\beta,c,d)<x<(\omega(L_\beta)-d)/c.$$

Fix again $\beta_0\in\mathbb{R}$ but such that $|\beta_0|>2$ and choose the alternatives

$$\beta_n=\beta_n(\vartheta),\qquad c_n=c_n(\xi),\qquad d_n=d_n(\eta)$$

of β_0, $c_0 = 1$, $d_0 = 0$ as

$$\begin{aligned}(\beta_0 - \beta_n)/\beta_0 &= \vartheta k^{-1/2}/\log(n/k),\\ 1 - c_n &= \xi k^{-1/2}/\beta_0,\\ d_n &= -\eta k^{-1/2}(n/k)^{1/\beta_0}/|\beta_0|\end{aligned}$$

where $\vartheta, \xi, \eta \in \mathbb{R}$. Observe that the alternative location parameter d_n is of smaller order than $k^{-1/2}$ iff the hypothetical shape parameter β_0 is negative.

For a family $F_{\beta,c,d}(x) = L_{\beta,c,d}(x)$, $x_0(\beta, c, d) < x < (\omega(L_\beta) - d)/c$ with $\limsup_{n\to\infty} |x_0(\beta_n, c_n, d_n)|^{\beta_0}(k/n) < 1$, where $k \to \infty$, $k/n \to 0$ as $n \to \infty$, we have the following result.

THEOREM 3.3. (LAN) *For $\beta_0 \in \mathbb{R}$ with $|\beta_0| > 2$ and $\beta_n = \beta_n(\vartheta)$, $c_n = c_n(\xi)$, $d_n = d_n(\eta)$ we have the expansion*

$$\begin{aligned}&\log\{d\mathcal{L}_{\beta_n,c_n,d_n}((Z_{n-i+1:n})_{i=k}^1)/d\mathcal{L}_{\beta_0,1,0}((Z_{n-i+1:n})_{i=k}^1)\}(Z_{n-i+1:n})_{i=k}^1\\ &\quad= (\vartheta + \xi + \eta)k^{1/2}(\beta_0 \log(|Z_{n-k+1:n}|) - \log(n/k)) - (\vartheta + \xi + \eta)^2/2\\ &\quad\quad+ \eta k^{-1/2}\sum_{i=1}^{k}\left(\frac{\beta_0 + 1}{\beta_0}\frac{1}{Z_{n-i+1:n}/Z_{n-k+1:n}} - 1\right)\\ &\quad\quad- \frac{\eta^2}{2\beta_0(\beta_0 + 2)} + o_{P_{\beta_0,1,0}}(1)\\ &\quad\underset{\mathcal{D}_{\beta_0,1,0}}{\longrightarrow} N(-(\vartheta + \xi + \eta)^2/2 - \eta^2/(2\beta_0(\beta_0 + 2)),\\ &\qquad\qquad (\vartheta + \xi + \eta)^2 + \eta^2/(\beta_0(\beta_0 + 2))).\end{aligned}$$

Note that under L_{β_0} the kth-largest order statistic $Z_{n-k+1:n}$ and the vector $(Z_{n-i+1:n}/Z_{n-k+1:n})_{i=k}^1$ are stochastically independent. The distribution of $(Z_{n-i+1:n}/Z_{n-k+1:n})_{i=k-1}^1$ equals further that of $(|V_{i:k-1}|)_{i=1}^{k-1}$, where $V_1, \dots, V_{k-1}$ are iid with common df L_{β_0}. This follows from Corollary 1.6.12 and Theorem 1.8.1 in Reiss (1989).

The central sequence in the preceding result is therefore the sum of two asymptotically independent terms based on $Z_{n-k+1:n}$ *and* the vector $(Z_{n-i+1:n}/Z_{n-k+1:n})_{i=k}^1$, with the vector $(Z_{n-i+1:n}/Z_{n-k+1:n})_{i=k}^1$ carrying information *only* about the location parameter d, and $Z_{n-k+1:n}$ containing *all* the information about the underlying shape and scale parameters β and c and a part of that about d.

The regularity condition $|\beta_0| > 2$ is crucial in various parts of the proof of Theorem 3.3. In particular it ensures that the second moments of $1/V_i$ are finite and therefore, the central limit theorem together with the preceding considerations imply that

$$k^{-1/2}(\beta_0(\beta_0 + 2))^{1/2}\sum_{i=1}^{k}\left(\frac{\beta_0 + 1}{\beta_0}\frac{1}{Z_{n-i+1:n}/Z_{n-k+1:n}} - 1\right) \underset{\mathcal{D}_{\beta_0,1,0}}{\longrightarrow} N(0, 1).$$

The limiting normal distribution in Theorem 3.3 is then a simple consequence of normal convolution. Recall that $k^{1/2}(\beta_0 \log(|Z_{n-k+1:n}|) - \log(n/k)) \underset{\mathcal{D}_0}{\rightarrow} N(0,1)$ as well. For a further discussion of the regularity condition $|\beta_0| > 2$ we refer to Hosking and Wallis (1987).

PROOF OF THEOREM 3.3. First note that if $|\beta_0| > 2$

$$(3.7)\quad P_{\beta_0,1,0}\{Z_{n-i+1:n} \in (x_0(\beta_n, c_n, d_n), (\omega(L_{\beta_0}) - d_n)/c_n), 1 \le i \le k\} \to_{n\to\infty} 1.$$

By (3.7) we can suppose for the rest of the proof that $Z_{n-i+1:n} \in (x_0(\beta_n, c_n, d_n), (\omega(L_{\beta_0}) - d_n)/c_n)$, $1 \le i \le k$, with underlying df L_{β_0}. Put again $Y_{n-k+1:n} := \beta_0 \log(|Z_{n-k+1:n}|)$ and $X_{(k)} := k^{1/2}(Y_{n-k+1:n} - \log(n/k))$. Then we have

$$\begin{aligned}(3.8)\quad & (L_{\beta_n,c_n,d_n}(Z_{n-k+1:n}) - L_{\beta_0,1,0}(Z_{n-k+1:n}))/L_{\beta_0,1,0}(Z_{n-k+1:n}) \\ &= (|Z_{n-k+1:n}|^{-\beta_0} - |c_n Z_{n-k+1:n} + d_n|^{-\beta_n})/(1 - |Z_{n-k+1:n}|^{-\beta_0}) \\ &= \frac{-k\exp(-X_{(k)}/k^{1/2})}{n - k\exp(-X_{(k)}/k^{1/2})} \\ &\quad \cdot (\exp\{-\beta_n \log(|c_n Z_{n-k+1:n} + d_n|) + \beta_0 \log(|Z_{n-k+1:n}|)\} - 1) \\ &= O_{P_{\beta_0,1,0}}(k^{1/2}/n),\end{aligned}$$

by Taylor expansion of exp at 0 and of log at 1, the definitions of β_n, c_n, d_n and the facts that $Z_{n-k+1:n}/(n/k)^{1/\beta_0} \to_{n\to\infty} 1$ and $Y_{n-k+1:n}/\log(n/k) \to_{n\to\infty} 1$ in $P_{\beta_0,1,0}$-probability. Equally, we have with $l_{\beta,c,d}$ denoting the density of $L_{\beta,c,d}$

$$\begin{aligned}(3.9)\quad & \sum_{i=1}^{k} \log\left\{\frac{l_{\beta_n,c_n,d_n}(Z_{n-i+1:n})}{l_{\beta_0,1,0}(Z_{n-i+1:n})}\right\} \\ &= \sum_{i=1}^{k}\left\{\log\left(\frac{\beta_n}{\beta_0}\right) - (\beta_n+1)\log\left(\left|1 + \frac{d_n}{c_n Z_{n-i+1:n}}\right|\right)\right. \\ &\qquad \left. - \beta_n \log(c_n) + \frac{\beta_0 - \beta_n}{\beta_0} Y_{n-i+1:n}\right\}.\end{aligned}$$

By repeating the arguments in the proof of Theorem 3.2, we obtain from (3.8) and (3.9) the expansion

$$\begin{aligned}(3.10)\quad & \log\{d\mathcal{L}_{\beta_n,c_n,d_n}((Z_{n-i+1:n})_{i=k}^{1})/d\mathcal{L}_{\beta_0,1,0}((Z_{n-i+1:n})_{i=k}^{1})\}(Z_{n-i+1:n})_{i=k}^{1} \\ &= \sum_{i=1}^{k}\left(\log\left\{\frac{l_{\beta_n,c_n,d_n}(Z_{n-i+1:n})}{l_{\beta_0,1,0}(Z_{n-i+1:n})}\right\} - \left(\frac{\beta_0-\beta_n}{\beta_0} Y_{n-k+1:n}\right.\right. \\ &\qquad \left.\left. - \beta_n \log(c_n) - \beta_n \log\left(\left|1 + \frac{d_n}{c_n Z_{n-k+1:n}}\right|\right)\right)\right) \\ &\quad + (n-k)\frac{L_{\beta_n,c_n,d_n}(Z_{n-k+1:n}) - L_{\beta_0,1,0}(Z_{n-k+1:n})}{L_{\beta_0,1,0}(Z_{n-k+1:n})}\end{aligned}$$

$$
\begin{aligned}
&+ k\Big(\frac{\beta_0-\beta_n}{\beta_0}Y_{n-k+1:n} - \beta_n\log(c_n) \\
&\qquad - \beta_n \log\Big(\Big|1+\frac{d_n}{c_nZ_{n-k+1:n}}\Big|\Big)\Big) \\
&+ o_{P_{\beta_0,1,0}}(1) \\
&= \beta_n \sum_{i=k}^{k}\Big(\log\Big(\Big|1+\frac{d_n}{c_nZ_{n-k+1:n}}\Big|\Big) \\
&\qquad - \frac{\beta_n+1}{\beta_n}\log\Big(\Big|1+\frac{d_n}{c_nZ_{n-i+1:n}}\Big|\Big)\Big) \\
&+ (\vartheta+\xi+\eta)k^{1/2}(X_{(k)}-\log(n/k)) \\
&- (\vartheta+\xi+\eta)^2/2 + o_{P_{\beta_0,1,0}}(1).
\end{aligned}
$$

The following expansion can be shown by conditioning on $Z_{n-k+1:n}=u$, in which case the (conditional) distribution of $(Z_{n-i+1:n})_{i=k-1}^{1}$ equals that of $(|u|V_{i:k-1})_{i=1}^{k-1}$, where $V_1, V_2, \dots$ are iid with common df L_{β_0}:

$$
\begin{aligned}
(3.11)\quad &\beta_n\sum_{i=1}^{k}\Big\{\log\Big(\Big|1+\frac{d_n}{c_nZ_{n-k+1:n}}\Big|\Big) - \frac{\beta_n+1}{\beta_n}\log\Big(\Big|1+\frac{d_n}{c_nZ_{n-i+1:n}}\Big|\Big)\Big\} \\
&= \eta k^{-1/2}\sum_{i=1}^{k}\Big\{\frac{\beta_0+1}{\beta_0}\frac{1}{Z_{n-i+1:n}/Z_{n-k+1:n}} - 1\Big\} \\
&\quad - \frac{\eta^2}{2\beta_0(\beta_0+2)} + o_{P_{\beta_0,1,0}}(1)
\end{aligned}
$$

as $Z_{n-k+1:n}/(n/k)^{1/\beta_0} \to_{n\to\infty} 1$ in $L_{\beta_0,1,0}$-probability. The assertion of Theorem 3.3 is now a consequence of formulas (3.10) and (3.11). □

We presently do not know, whether LAN of extreme order statistics can be established for underlying dfs, which do *not* belong to a δ-neighborhood of a GPD such as a normal df. Various examples, which we have computed, give rise to the conjecture that this is actually not possible.

Acknowledgements

The author is grateful to Frank Marohn for helpful discussions and the referees for their constructive criticism from which the paper has benefited a lot.

References

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

Castillo, E., Galambos, J. and Sarabia, J. M. (1989). The selection of the domain of attraction of an extreme value distribution from a set of data, *Extreme Value Theory* (eds. J. Hüsler and R.-D. Reiss), *Lecture Notes in Statist.*, **51**, 181–190, Springer, New York.

Csörgő, S. and Mason, D. M. (1985). Central limit theorems for sums of extreme values, *Mathematical Proceedings of the Cambridge Philosophical Society*, **98**, 547–558.

Dekkers, A. L. M. and de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation, *Ann. Statist.*, **17**, 1795–1832.

Falk, M. (1992). On testing the extreme value index via the POT-method, *Ann. Statist.* (to appear).

Falk, M., Hüsler, J. and Reiss, R.-D. (1994). *Laws of Small Numbers: Extremes and Rare Events*, DMV Seminar 23, Birkhäuser, Basel.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics* (2nd ed.), Krieger, Malabar.

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

Gumbel, E. J. (1958). *Statistics of Extremes*, Columbia University Press, New York.

Hall, P. and Welsh, A. H. (1985). Adaptive estimates of parameters of regular variation, *Ann. Statist.*, **13**, 331–341.

Hasofer, A. M. and Wang, Z. (1992). A test for extreme value domain of attraction, *J. Amer. Statist. Assoc.*, **87**, 171–177.

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

Hosking, J. R. M. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, **29**, 339–349.

Janssen, A. and Marohn, F. (1994). On statistical information of extreme order statistics, local extreme value alternatives, and Poisson point processes, *J. Multivariate Anal.*, **48**, 1–30.

Janssen, A. and Reiss, R.-D. (1988). Comparison of location models of Weibull type samples and extreme value processes, *Probab. Theory Related Fields*, **78**, 273–292.

LeCam, L. (1960). Locally asymptotically normal families of distributions, *University of California Publications in Statistics*, **3**, 37–98.

LeCam, L. (1986). Asymptotic methods in statistical decision theory, *Springer Ser. Statist.*, Springer, New York.

LeCam, L. and Yang, G. L. (1990). Asymptotics in statistics (some basic concepts), *Springer Ser. Statist.*, Springer, New York.

Marohn, F. (1991). Global sufficiency of extreme order statistics in location models of Weibull type, *Probab. Theory Related Fields*, **88**, 261–268.

Marohn, F. (1994*a*). On testing the exponential and Gumbel distribution, *Extreme Value Theory and Applications* (eds. J. Galambos, J. Lechner and E. Simiu), 159–174, Kluwer, Dordrecht.

Marohn, F. (1994*b*). Asymptotic sufficiency of order statistics for almost regular Weibull type densities, *Statist. Decisions*, **12**, 385–393.

Marohn, F. (1995). Neglecting observations in Gaussian sequences of statistical experiments, *Statist. Decisions*, **13**, 83–92.

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

Reiss, R.-D. (1989). Approximate distributions of order statistics (with applications to nonparametric statistics), *Springer Ser. Statist.*, Springer, New York.

Reiss, R.-D. (1993). A course on point processes, *Springer Ser. Statist.*, Springer, New York.

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

Strasser, H. (1985). Mathematical theory of statistics, *De Gruyter Stud. Math.*, **7**, De Gruyter, Berlin.

Wei, X. (1992). Asymptotically efficient estimation of the index of regular variation, Ph.D. thesis, Department of Statistics, University of Michigan.