The Annals of Statistics
1995, Vol. 23, No. 6, 2013–2035

ON TESTING THE EXTREME VALUE INDEX VIA THE POT-METHOD

BY MICHAEL FALK

Katholische Universität Eichstätt

Consider an iid sample $Y_1, \dots, Y_n$ of random variables with common distribution function F, whose upper tail belongs to a neighborhood of the upper tail of a generalized Pareto distribution H_β, $\beta \in \mathbb{R}$. We investigate the testing problem $\beta = \beta_0$ against a sequence $\beta = \beta_n$ of contiguous alternatives, based on the point processes N_n of the exceedances among Y_i over a sequence of thresholds t_n. It turns out that the (random) number of exceedances $\tau(n)$ over t_n is the central sequence for the log-likelihood ratio $d\mathscr{L}_{\beta_n}(N_n)/d\mathscr{L}_{\beta_0}(N_n)$, yielding its local asymptotic normality (LAN). This result implies in particular that $\tau(n)$ carries asymptotically all the information about the underlying parameter β, which is contained in N_n. We establish sharp bounds for the rate at which $\tau(n)$ becomes asymptotically sufficient, which show, however, that this is quite a poor rate. These results remain true if we add an unknown scale parameter.

0. Introduction. Consider a distribution function (d.f.) F on the real line whose upper tail belongs to a member of a parametric family of d.f.'s. To be precise, we suppose that there exists an unknown root $x_0 \in \mathbb{R}$ such that

$$F(x) = F_\beta(x) \quad \text{for } x \ge x_0, \tag{M}$$

where $\{F_\beta : \beta \in \Theta\}$ is a family of d.f.'s, parametrized by the elements β from some parameter space Θ. Our problem is to deduce statistical inference about the unknown parameter β from an iid sample $Y_1, \dots, Y_n$ with common d.f. F.

A model assumption on the upper tail of the underlying d.f. such as (M) is, for example, indispensable if one is interested in *extreme quantities* of the underlying d.f., that is, such quantities of F which are usually outside the range of the given data $Y_1, \dots, Y_n$. A typical example is inference about extreme quantiles $F^{-1}(p) := \inf\{t \in \mathbb{R}: F(t) \ge p\}$ with p being close to 1. This problem is typically tackled by hydrologists for the prediction of large floods [cf. Hosking and Wallis (1987) and the literature cited therein].

Statistical inference about the parameter β in model (M) can clearly be deduced only from those observations among $Y_1, \dots, Y_n$ which are *large* in some sense. There are two obvious but different ways to define an observation

Received November 1992; revised November 1994.

AMS 1991 *subject classifications*. Primary 62F05; secondary 60G70, 60G55.

Key words and phrases. Generalized Pareto distribution, δ-neighborhood, peaks over threshold, point process of exceedances, log-likelihood ratio, local asymptotic normality, asymptotic sufficiency, Hellinger distance.

Y_i as large:

(a) if it is among the k largest *order statistics* $Y_{n-k+1:n}, Y_{n-k+2:n}, \dots, Y_{n:n}$, where $Y_{1:n} \le \cdots \le Y_{n:n}$ denote the ordered values of $Y_1, \dots, Y_n$;

(b) if it exceeds a given high *threshold* t_n, say.

While (a) has been the center of interest of *extreme value theory* since its beginning [cf. de Haan (1970) and Galambos (1987)], approach (b) has only recently found increasing interest [cf. Smith (1987), Davison and Smith (1990), Falk, Hüsler and Reiss (1994) and the references given therein]. Although it seems to be more natural to call a value large if it exceeds some specific threshold, the minor attention paid to approach (b) might be due to the fact that this method generates a vector $(V_1, \dots, V_{\tau(n)})$ of random length $\tau(n)$, say, if we apply it to $Y_1, \dots, Y_n$. Here $V_1, V_2, \dots, V_{\tau(n)}$ denote those values Y_i among $Y_1, \dots, Y_n$ which exceed the threshold t_n, arranged in the order of their outcome; their total number $\tau(n)$ is then binomially distributed $B(n, 1 - F(t_n))$. This *peaks-over-threshold* (POT) method suffers therefore from a dimensionality problem, and the formulation of convergence results is not obvious.

In recent years, however, the theory of *point processes* has become more and more important in different statistical fields, due to the fact that it provides a way to analyze data in a dimension-free way. This quite general concept is therefore tailor-made for approach (b). While the recent book by Falk, Hüsler and Reiss (1994) focuses on applications of the point process approach to extreme value models, excellent introductions to general point process theory are provided by the monographs by Daley and Vere-Jones (1988) and Reiss (1993). We briefly summarize the very few elements from the general theory which we need for dealing with our particular model (M).

First of all, we identify a point $x \in \mathbb{R}$ with the pertaining Dirac measure $\varepsilon_x(B) = 1_B(x) = 1$ if $x \in B$ and 0 otherwise. We thus identify the *excess* $V_i - t_n$ with the random Dirac measure

$$\varepsilon_{V_i - t_n}(B) = 1_B(V_i - t_n), \qquad B \in \mathscr{B},$$

on the Borel σ-field $\mathscr{B}$ of $\mathbb{R}$. We prefer to work with the excesses $V_i - t_n$ rather than with the exceedances V_i over t_n themselves, as their range $(0, \infty)$ is kept fixed by this shift.

A mathematically precise description of the vector $(V_1 - t_n, \dots, V_{\tau(n)} - t_n)$ of excesses is then

$$N_n(B) := \sum_{i=1}^{n} \varepsilon_{Y_i - t_n}(B)\, \varepsilon_{Y_i}(t_n, \infty) = \sum_{i=1}^{\tau(n)} \varepsilon_{V_i - t_n}(B), \qquad B \in \mathscr{B},$$

where $\tau(n) := N_n(\mathbb{R}) = N_n((0, \infty))$. Note that N_n is a random element in the set $\mathbb{M} := \{\mu = \sum_{i=1}^{n} \varepsilon_{x_i}\colon x_1, \dots, x_n \in \mathbb{R},\ n = 0, 1, 2, \dots\}$ of (finite) point measures on $(\mathbb{R}, \mathscr{B})$, equipped with the smallest σ-field $\mathscr{M}$ such that for any $A \in \mathscr{B}$ the projection $\pi_A\colon \mathbb{M} \to \{0, 1 \dots\}$, $\pi_A(\mu) := \mu(A)$, is measurable. As such, N_n is called a *point process*. [For technical details we refer to Reiss (1993), Section 1.1.]

The following lemma is crucial for the POT-method; it is a special case of a general result for truncated empirical processes [cf. Reiss (1993), Theorem 1.4.1]. We let $\mathscr{L}(N_n)$ denote the distribution of N_n.

LEMMA 0.1. *Let $Y_1, Y_2, \dots$ be independent copies of a random variable (r.v.) Y with d.f. F, and we choose $t_n \in \mathbb{R}$ such that $0 < 1 - F(t_n) < 1$. Then*

$$\mathscr{L}(N_n) = \mathscr{L}\left(\sum_{i=1}^{\tau(n)} \varepsilon_{V_i - t_n}\right) = \mathscr{L}\left(\sum_{i=1}^{\tau(n)} \varepsilon_{U_i}\right),$$

where $\tau(n)$ is $B(n, 1 - F(t_n))$ distributed, $U_1, U_2, \dots, U_n$ are iid r.v.'s with common d.f.

$$(1)\quad B(x) := P(Y - t_n \le x | Y - t_n > 0) = 1 - \frac{1 - F(t_n + x)}{1 - F(t_n)}, \qquad x \ge 0,$$

and $\tau(n)$ and the vector $(U_1, \dots, U_n)$ are independent.

The preceding result provides a rather easy access to the investigation of $N_n = \sum_{i=1}^{\tau(n)} \varepsilon_{V_1 - t_n}$ by decomposing it into two *independent* parts, namely, the *values* of the excesses $V_i - t_n$ and their *number* $\tau(n)$. The excesses $V_i - t_n$ are independent with common d.f. $B(\cdot)$ as given in the preceding result, and $\tau(n)$ is $B(n, 1 - F(t_n))$, distributed.

We will investigate, within the model (M), the testing problem

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

converging to β_0, based on the POT-method with increasing thresholds t_n for particular families F_β of d.f.'s, which we will introduce in the sequel.

According to the results in Balkema and de Haan (1974) or in Rychlik [(1992), Theorem 2.1] [see also Falk, Hüsler and Reiss (1994), Theorem 1.3.5], the only set of possible nondegenerate weak limits of the excess d.f. $B(x) = 1 - (1 - F(t_n + x))/(1 - F(t_n))$, $x \ge 0$, as $F(t_n) \to 1$, is under mild regularity conditions on the sequence t_n the class of *generalized Pareto d.f.'s* (GPD's). In their von Mises parametrization, these GPD's are for $\beta \in \mathbb{R}$ defined by

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

Interpret H_0 as $H_0(x) = \lim_{\beta \to 0} H_\beta(x) = \exp(-x)$, $x \ge 0$. Note that H_β has support $(0, -1/\beta)$ if $\beta < 0$ and $(0, \infty)$ if $\beta \ge 0$.

This result makes a GPD a natural model for the excess d.f. B in (1), in which case the upper tail of F equals that of a shifted GPD. This is, for example, a quite common approach in insurance mathematics to model large claims [cf. Teugels (1984)] or in hydrology to model large floods [cf. Hosking and Wallis (1987)]. We will therefore consider in this paper such d.f.'s F, whose upper tails are in certain neighborhoods of that of a GPD.

Denote by $\omega(G) := \sup\{x \in \mathbb{R}: G(x) < 1\}$ the right endpoint of the support of a d.f. G. The upper tail of a d.f. F belongs to the *δ-neighborhood* of some GPD H_β if $\omega(F) = \omega(H_\beta)$ and F has density f for some (commonly unknown) $x_0 < \omega(F)$ such that, for some constant $C > 0$,

$$|f(x)/h_\beta(x) - 1| \leq C\big(1 - H_\beta(x)\big)^\delta, \qquad x \geq x_0,$$

where h_β denotes the density of H_β. The importance of δ-neighborhoods of GPD's in extreme value theory, their derivation from von Mises conditions, their connection with rates of convergence of extremes and so on is extensively described in Falk, Hüsler and Reiss [(1994), Section 1.3 and Chapter 2].

We will therefore consider for the testing problem $\beta = \beta_0$ against $\beta_n \neq \beta_0$ in this paper a parametric family $\{F_\beta\colon \beta \in \mathbb{R}\}$ such that for each $\beta \in \mathbb{R}$ the d.f. F_β is in a δ-neighborhood of a GPD H_β.

While *estimation* of the extreme value index β has been extensively studied in the literature [see, e.g., Smith (1987); Reiss (1989), Chapter 9; Dekkers and de Haan (1989); Falk, Hüsler and Reiss (1994), Sections 2.4 and 2.5; and the literature cited therein], comparatively little has been published on testing of β, in particular, on testing $\beta = 0$ against a sequence $\beta_n \neq 0$ [Castillo, Galambos and Sarabia (1989), Gomes (1989) and Hasofer and Wang (1992)]. In particular, the powerful theory of *local asymptotic normality* (LAN) of statistical experiments, developed by Le Cam [cf. Le Cam (1960, 1986), Le Cam and Yang (1990) and Strasser (1985)], has been applied to extreme value problems as yet in only a few and quite recent papers [Marohn (1991, 1994a, b), Janssen and Marohn (1994) and Wei (1992); related papers are Janssen and Reiss (1988), Höpfner and Jacod (1993) and Höpfner (1994)].

In the present paper we will test $\beta = \beta_0$ against a sequence β_n of contiguous alternatives, where the tests are based on the point processes N_n of excesses. This will be done within Le Cam's concept of local asymptotic normality. We index expectations E_ϑ, distributions $\mathscr{L}_\vartheta$ and so on by the underlying parameter. An expansion of the log-likelihood ratio

$$L_n = \log\left\{\frac{d\mathscr{L}_{\beta_n}(N_n)}{d\mathscr{L}_{\beta_0}(N_n)}\right\}(N_n)$$

reveals that $\tau(n)$ is the *central sequence* for our testing problem β_n against β_0, yielding that L_n is under β_0 local asymptotically normal distributed, LAN for short. This implies that asymptotically $\tau(n)$ is sufficient for β_n against β_0; that is, the complete information contained in N_n about the structural parameter β is already contained in the number $\tau(n)$ of exceedances only. Asymptotically optimal tests for $\beta = \beta_0$ against β_n can therefore solely be based on τ_n, and standard results from LAN theory provide the asymptotic power function as well; see the remarks after Theorem 1.1. For a precise definition of asymptotic sufficiency and its connection with central sequences we refer to Le Cam and Yang [(1990), Proposition 2 in Section 5.3].

We establish a bound for this increasing sufficiency of $\tau(n)$ by establishing a sharp bound for the *Hellinger distance* $H(\mathscr{L}_\beta(N_n), \mathscr{L}_\beta(N_n^*))$ between the distribution of $N_n = \sum_{i=1}^{\tau(n)} \varepsilon_{V_i - t_n}$ and that of

$$N_n^* := \sum_{i=1}^{\tau(n)} \varepsilon_{W_i}$$

under $\beta = \beta_n$, where $\tau(n) = N_n((0, \infty))$ is kept, but $W_1, W_2, \dots$ are excesses based on the hypothetical F_{β_0}. Under $\beta = \beta_0$ the distributions of N_n and N_n^* coincide. This bound converges to zero as $\beta_n \to_{n\to\infty} \beta_0$, but at a rather slow rate. These considerations are the content of Section 1.

Adding an unknown scale parameter $c > 0$ to F_β and considering $F_\beta(cx)$, one might guess that the excesses contribute to the information about the scale parameter c, when testing

$$(\beta_n, c_n) \quad \text{against} \quad (\beta_0, c_0).$$

However, the results of Section 1 carry over to this problem; that is, $\tau(n)$ is still asymptotically sufficient, which is shown in Section 2.

Various examples which we have computed indicate that the log-likelihood ratio L_n is no longer LAN if F_{β_0} and F_{β_n} are not in δ-neighborhoods of GPD's. The proof of the conjecture that L_n is actually LAN *only if* F_{β_0} and F_{β_n} are in δ-neighborhoods of GPD's is an open problem.

1. Testing the extreme value index. Consider at first a d.f. F whose upper tail coincides with that of a GPD in the sense of model (M), that is, $F(x) = H_\beta(x)$, $x \geq x_0$, for some unknown $x_0 = x_0(\beta)$ and $\beta \in \mathbb{R}$. The case $F(x) = H_\beta(cx)$ for some (unknown) scale parameter $c > 0$ is considered in the next section. The d.f. of the excess distribution $B_\beta(x) = P(V - t_n \leq x) = P(Y \leq t_n + x | Y > t_n)$, $x \geq 0$, is, by Lemma 0.1,

$$\begin{aligned} B_\beta(x) &= 1 - \frac{1 - H_\beta(x + t_n)}{1 - H_\beta(t_n)} \\ &= 1 - \left(1 + \beta \frac{x}{1 + \beta t_n}\right)^{-1/\beta}, \qquad 0 \leq x < \omega(H_\beta) - t_n, \end{aligned}$$

with $B_0(x) = \lim_{\beta \to 0} B_\beta(x) = 1 - \exp(-x)$, $x \geq 0$, if $0 < H_\beta(t_n) < 1$ and $t_n > x_0$. The excess distribution B_β has density

$$b_\beta(x) = \frac{1}{1 + \beta t_n}\left(1 + \beta \frac{x}{1 + \beta t_n}\right)^{-1/\beta - 1}, \qquad 0 \leq x < \omega(H_\beta) - t_n,$$

with $b_0(x) = \lim_{\beta \to 0} b_\beta(x) = \exp(-x)$, $x \geq 0$. First we consider the hypothesis $\beta_0 = 0$. Choose the threshold

$$t_n := \log(na_n)$$

and the alternatives

$$\beta_n := \beta_n(\vartheta) := 2\vartheta a_n^{1/2}/t_n^2, \qquad \vartheta \in \mathbb{R},$$

where the sequence $a_n > 0$, $n \in \mathbb{N}$, satisfies $a_n \to 0$, $na_n \to \infty$ as $n \to \infty$. Notice that the threshold t_n may converge to infinity at any prescribed rate below $\log(n)$ by a suitable choice of a_n, yielding an asymptotically increasing number of expected excesses. However, once t_n has been chosen, the preceding definition of the alternatives β_n is required for a nondegenerate normal limit of the log-likelihood ratio in Theorem 1.1. The same applies to subsequent results.

The following result shows that the number $\tau(n)$ of exceedances is the central sequence for testing $\beta = 0$ against $\beta = \beta_n$; that is, Theorem 1.1 reveals that the complete information about the underlying extreme value index β contained in N_n is asymptotically already contained in the number $\tau(n) = N_n((0, \infty))$ of exceedances of the original data over the threshold t_n. By $\to_{\mathscr{D}_\beta}$ we denote weak convergence under the parameter β, and by $o_{P_\beta}(1)$ a stochastic remainder term which converges to zero in probability under β as the sample size n increases.

THEOREM 1.1 (LAN). *Under the hypothesis $\beta = 0$ we have, for any $\vartheta \in \mathbb{R}$,*

$$\log\left\{\frac{d\mathscr{L}_{\beta_n}(N_n)}{d\mathscr{L}_0(N_n)}\right\}(N_n)$$
$$= \vartheta a_n^{1/2}\left(\tau(n) - a_n^{-1}\right) - \frac{\vartheta^2}{2} + o_{P_0}(1) \to_{\mathscr{D}_0} N\left(-\frac{\vartheta^2}{2}, \vartheta^2\right).$$

Notice that the ad hoc test statistic for testing β_n against β_0 based on N_n is

$$T_n := \sum_{i=1}^{\tau(n)} \log\left\{\frac{b_{\beta_n}}{b_{\beta_0}}(V_i - t_n)\right\},$$

which is suggested by the Neyman–Pearson lemma, with fixed sample size n replaced by the (independent) r.v. $\tau(n)$. The proof of Theorem 1.1 shows, however, that T_n is of order $o_P(1)$ under β_n and $\beta_0 = 0$. Although T_n seems to be a natural and powerful test statistic for testing β_n against $\beta_0 = 0$, it is therefore not adequate, as it cannot distinguish asymptotically between β_n and 0, but $(\tau(n) - a_n^{-1})a_n^{1/2}$ does. Theorem 1.3, as well as Theorem 2.1, shows that this remains true for a hypothetical value $\beta_0 \neq 0$, even if an unknown scale parameter is added.

The preceding result enables us to apply the powerful general LAN theory to our particular testing problem [cf. Le Cam (1986), Le Cam and Yang (1990), and Strasser (1985)]. Le Cam's first and third lemmas imply, in particular, asymptotic equivalence between the log-likelihood ratio $\log\{d\mathscr{L}_{\beta_n}(N_n)/d\mathscr{L}_0(N_n)\}(N_n)$ and $\vartheta a_n^{1/2}(\tau(n) - a_n^{-1}) - \vartheta^2/2$ also under the *contiguous* alternatives F_{β_n}; that is, Theorem 1.1 implies

$$\log\left\{\frac{d\mathscr{L}_{\beta_n}(N_n)}{d\mathscr{L}_0(N_n)}\right\}(N_n) = \vartheta a_n^{1/2}\left(\tau(n) - a_n^{-1}\right) - \frac{\vartheta^2}{2} + o_{P_{\beta_n}}(1)$$

$$\to_{\mathscr{D}_{\beta_n}} N\left(\frac{\vartheta^2}{2}, \vartheta^2\right).$$

Denote by Φ the standard normal d.f., and set $u_\alpha := \Phi^{-1}(1-\alpha)$, $0<\alpha<1$. Then, by the preceding result and the Neyman–Pearson lemma,

$$\varphi_n(N_n) := 1_{(u_\alpha,\infty)}\left(\operatorname{sign}(\vartheta) a_n^{1/2}\left(\tau(n) - a_n^{-1}\right)\right) \tag{2}$$

is an asymptotically optimal level α test, based on N_n, for testing $\beta = 0$ against $\beta_n = \beta_n(\vartheta)$, with asymptotic power function

$$\psi(\vartheta) := \lim_{n\to\infty} E_{\beta_n(\vartheta)}(\varphi_n) = 1 - \Phi(u_\alpha - |\vartheta|), \qquad \vartheta \in \mathbb{R}. \tag{3}$$

Note that $\varphi_n(N_n)$ is asymptotically optimal uniformly for $\vartheta > 0$ and $\vartheta < 0$.

PROOF OF THEOREM 1.1. From Reiss [(1993), Example 3.1.2] we conclude that $\mathscr{L}_{\beta_n}(N_n)$ has $\mathscr{L}_0(N_n)$-density

$$g(\mu) = \left(\prod_{i=1}^{\mu(\mathbb{R}_+)} \frac{b_{\beta_n}(x_i)}{b_0(x_i)}\right)\left(\frac{1 - H_{\beta_n}(t_n)}{1 - H_0(t_n)}\right)^{\mu(\mathbb{R}_+)}\left(\frac{H_{\beta_n}(t_n)}{H_0(t_n)}\right)^{n-\mu(\mathbb{R}_+)}$$

if $\mu = \sum_{i=1}^{\mu(\mathbb{R}_+)} \varepsilon_{x_i}$ and $0 \le \mu(\mathbb{R}_+) \le n$. Consequently,

$$\begin{aligned}\log\left\{\frac{d\mathscr{L}_{\beta_n}(N_n)}{d\mathscr{L}_0(N_n)}\right\}(\mu) &= \int \log\left\{\frac{b_{\beta_n}(y)}{b_0(y)}\right\}\mu(dy)\\ &\quad + \mu(\mathbb{R}_+)\log\left\{\frac{1 - H_{\beta_n}(t_n)}{1 - H_0(t_n)}\right\}\\ &\quad + (n - \mu(\mathbb{R}_+))\log\left(\frac{H_\beta(t_n)}{H_0(t_n)}\right).\end{aligned}$$

In the following we will drop the index n of β_n for the sake of a clear presentation. Recall that $na_n = \exp(t_n)$, and observe that $0 < \log(na_n) < \omega(H_{\beta_n})$ if n is large. Taylor expansions of log at 1 and exp at 0 imply the following fact.

FACT 1. We have $H_\beta(t_n) - H_0(t_n) = -na_n^{-1}(\vartheta a_n^{1/2} + O(a_n))$.

Consequently,

$$\tau(n)\log\left\{\frac{1-H_\beta(t_n)}{1-H_0(t_n)}\right\}+(n-\tau(n))\log\left(\frac{H_\beta(t_n)}{H_0(t_n)}\right)$$

$$(4)\qquad =\tau(n)\left\{na_n\big(H_0(t_n)-H_\beta(t_n)\big)-\frac{\big(na_n(H_0(t_n)-H_\beta(t_n))\big)^2}{2}+O\big(a_n^{3/2}\big)\right\}$$
$$+(n-\tau(n))\left\{\big(1-(na_n)^{-1}\big)^{-1}\big(H_\beta(t_n)-H_0(t_n)\big)+O\left(\frac{a_n}{(na_n)^2}\right)\right\}.$$

Recall that under $\beta=0$ the r.v. $\tau(n)$ is $B(n,1-H_0(t_n))=B(n,1/(na_n))$-distributed and hence, by Fact 1,

$$\tau(n)\frac{\big(na_n(H_0(t_n)-H_\beta(t_n))\big)^2}{2}=\tau(n)\frac{\vartheta^2a_n(1+o(1))}{2}\to\frac{\vartheta^2}{2}$$

in probability, since $\tau(n)a_n$ converges to 1. Equally, $\tau(n)a_n^{3/2}$ and $(n-\tau(n))a_n/(na_n)^2$ both converge to 0. Finally, we order the remaining terms in (4) as follows:

$$\tau(n)na_n\big(H_0(t_n)-H_\beta(t_n)\big)+\frac{n-\tau(n)}{1-(na_n)^{-1}}\big(H_\beta(t_n)-H_0(t_n)\big)$$
$$=\big(\tau(n)-a_n^{-1}\big)na_n\big(H_0(t_n)-H_\beta(t_n)\big)$$
$$+\left\{n\left(\frac{1}{1-(na_n)^{-1}}-1\right)-\frac{\tau(n)}{1-(na_n)^{-1}}\right\}\big(H_\beta(t_n)-H_0(t_n)\big)$$
$$=\big(\tau(n)-a_n^{-1}\big)\vartheta a_n^{1/2}(1+o(1))+\frac{a_n^{-1}-\tau(n)}{1-(na_n)^{-1}}\big(H_\beta(t_n)-H_0(t_n)\big)$$
$$=\big(\tau(n)-a_n^{-1}\big)\vartheta a_n^{1/2}+o_{P_0}(1),$$

as $(\tau(n)-a_n^{-1})a_n^{1/2}$ is asymptotically standard normal and $H_\beta(t_n)-H_0(t_n)$ is, by Fact 1, of order $O(a_n^{1/2}/(na_n))=o(a_n^{1/2})$. Thus we have shown so far that

$$\tau(n)\log\left\{\frac{1-H_\beta(t_n)}{1-H_0(t_n)}\right\}+(n-\tau(n))\log\left(\frac{H_\beta(t_n)}{H_0(t_n)}\right)$$
$$=\big(\tau(n)-a_n^{-1}\big)\vartheta a_n^{1/2}-\frac{\vartheta^2}{2}+o_{P_0}(1).$$

In order to prove Theorem 1.1 it therefore remains to show that, under $\beta=0$,

$$(5)\qquad \int_0^\infty\log\left\{\frac{b_\beta(y)}{b_0(y)}\right\}N_n(dy)=o_{P_0}(1).$$

First observe that

$$\int_{1/|\beta|^{1/2}}^{\infty} \log\left\{\frac{b_\beta(y)}{b_0(y)}\right\} N_n(dy) = o_{P_0}(1).$$

By making use of Lemma 0.1 this is immediate from

$$\begin{aligned}
&P_0\left(\left|\int_{1/|\beta|^{-1/2}}^{\infty} \log\left\{\frac{b_\beta(y)}{b_0(y)}\right\} N_n(dy)\right| > \varepsilon\right)\\
&\quad = P_0\left(\left|\sum_{i=1}^{\tau(n)} \log\left\{\frac{b_\beta(U_i)}{b_0(U_i)}\right\} \varepsilon_{U_i}\left(\left[|\beta|^{-1/2}, \infty\right)\right)\right| > \varepsilon\right)\\
&\quad \le P_0\left(U_i \ge |\beta|^{-1/2} \text{ for some } i \in \{1, \dots, \tau(n)\}\right) \qquad (6)\\
&\quad = \sum_{m=1}^{n} P_0\left(U_i \ge |\beta|^{-1/2} \text{ for some } i \in \{1, \dots, m\}\right) P_0(\tau(n) = m)\\
&\quad \le \sum_{m=1}^{n} m P_0\left(U_i \ge |\beta|^{-1/2}\right) P_0(\tau(n) = m)\\
&\quad = \left(1 - H_0\left(|\beta|^{-1/2}\right)\right) E_0(\tau(n)) = o(1),
\end{aligned}$$

where $\varepsilon > 0$ is arbitrary. By using again the expansion $\log(1+\varepsilon) = \varepsilon - \varepsilon^2/2 + O(\varepsilon^3)$ for $\varepsilon \to 0$, we can write

$$\begin{aligned}
&\int_0^{1/|\beta|^{1/2}} \log\left\{\frac{b_\beta(y)}{b_0(y)}\right\} N_n(dy)\\
&\quad = \int_0^{1/|\beta|^{1/2}} y - \frac{1+\beta}{\beta} \log\left(1 + \frac{\beta y}{1+\beta t_n}\right) - \log(1+\beta t_n) N_n(dy)\\
&\quad = \int_0^{1/|\beta|^{1/2}} y - \frac{1+\beta}{\beta}\left\{\frac{\beta y}{1+\beta t_n} - \frac{\beta^2 y^2}{2(1+\beta t_n)^2} + O(\beta^3 y^3)\right\}\\
&\qquad - \beta t_n + O(\beta^2 t_n^2) N_n(dy)\\
&\quad = \int_0^{1/|\beta|^{1/2}} y - (1+\beta)\left\{\frac{y}{1+\beta t_n} - \frac{\beta y^2}{2(1+\beta t_n)^2} + O(\beta^2 y^3)\right\}\\
&\qquad - \beta t_n N_n(dy) + o_{P_0}(1),
\end{aligned}$$

as $\int_0^{1/|\beta|^{1/2}} \beta^2 t_n^2 N_n(dy) \le \beta^2 t_n^2 \tau(n) = 4\vartheta^2 a_n \tau(n)/t_n^2 = o_{P_0}(1)$. By the inequalities in (6) we can extend the preceding integral again over the range

$[0, \infty)$ and obtain that it is asymptotically equivalent to

$$\int_0^\infty \left(1 - \frac{1+\beta}{1+\beta t_n}\right) y + \frac{\beta}{2(1+\beta t_n)^2} y^2 - \beta t_n N_n(dy)$$

$$= \beta \sum_{i=1}^{\tau(n)} \left(\frac{t_n - 1}{1+\beta t_n}(U_i - 1) + \frac{1}{(1+\beta t_n)^2}\left(\frac{U_i^2}{2} - 1\right)\right)$$

$$+ \beta\tau(n)\left\{\frac{t_n - 1}{1+\beta t_n} + \frac{1}{(1+\beta t_n)^2} - t_n\right\}$$

$$= \left(\beta\tau(n)^{1/2}\frac{t_n-1}{1+\beta t_n}\right)\tau(n)^{-1/2}\sum_{i=1}^{\tau(n)}(U_i - 1)$$

$$+ \left(\frac{\beta\tau(n)^{1/2}}{(1+\beta t_n)^2}\right)\tau(n)^{-1/2}\sum_{i=1}^{\tau(n)}\left(\frac{U_i^2}{2} - 1\right) - \frac{\beta^2 t_n + \beta^2 t_n^2 + \beta^3 t_n^3}{(1+\beta t_n)^2}\tau(n)$$

$$= o_{P_0}(1),$$

by the definition of β, the central limit theorem and the fact that $\tau(n)a_n \to 1$ in probability. Recall that by Lemma 0.1 the r.v.'s $\tau(n)$ and $V_1, V_2, \ldots$ are independent. This implies (5) and completes the proof of Theorem 1.1. □

The preceding result remains true if we replace the parametric family $\{H_\beta: \beta \in \mathbb{R}\}$ of GPD's by d.f.'s F_β in the δ-neighborhood of F_β. Fix $C, \delta > 0$; again choose $\beta_n = \beta_n(\vartheta) = 2\vartheta a_n^{1/2}/\log^2(na_n)$, but this time consider, for $\beta = \beta_n$ and $\beta = 0$, d.f.'s F_β with $\omega(F_\beta) = \omega(H_\beta)$ having density f_β on $[x_0(\beta), \omega(F_\beta))$ such that

$$\left|\frac{f_\beta(x)}{h_\beta(x)} - 1\right| \le C\big(1 - H_\beta(x)\big)^\delta, \qquad x \in \big[x_0(\beta), \omega(F_\beta)\big).$$

If we now require that $x_0(\beta_n) \le t_n = \log(na_n)$, $n \in \mathbb{N}$, then it turns out that the number of exceedances over the threshold t_n is again the central sequence for testing $\beta = 0$ against $\beta = \beta_n(\vartheta)$ and is therefore asymptotically sufficient, under the additional assumption $a_n^{1/2}(na_n)^\delta \to \infty$ as $n \to \infty$. To be precise, we have again under $\beta = 0$ the expansion

$$\log\left\{\frac{d\mathscr{L}_{\beta_n}(N_n)}{d\mathscr{L}_0(N_n)}\right\}(N_n)$$

$$= \vartheta a_n^{1/2}\big(\tau(n) - E_0(\tau(n))\big) - \frac{\vartheta^2}{2} + o_{P_0}(1)$$

$$= \vartheta a_n^{1/2}\big(\tau(n) - a_n^{-1}\big) - \frac{\vartheta^2}{2} + o_{P_0}(1) \to_{\mathscr{D}_0} N\left(-\frac{\vartheta^2}{2}, \vartheta^2\right).$$

This can be shown along the lines of the proof of Theorem 1.1 by using Fact 2 in the proof of the following bound for the sufficiency of $\tau(n)$.

Consider again the empirical point process of the excesses

$$N_n = \sum_{i=1}^{\tau(n)} \varepsilon_{V_i - t_n}.$$

If $\tau(n)$ essentially contains all the information about β_n delivered by N_n, then the error should be small if we replace the actual data $V_1 - t_n, \dots, V_{\tau(n)} - t_n$ simply by ideal standard exponential r.v.'s $W_1, \dots, W_{\tau(n)}$, with the sequence $W_1, W_2, \dots$ being independent of $\tau(n)$. Define therefore

$$N_n^* := \sum_{i=1}^{\tau(n)} \varepsilon_{W_i},$$

which is generated from N_n by just replacing $V_i - t_n$ by W_i.

The following result estimates the error, when replacing N_n by N_n^*; it provides therefore a bound for the sufficiency of $\tau(n)$. By $H(\cdot\,,\cdot)$ we denote the *Hellinger distance* between the distributions of random elements on the same sample space. Precisely, let Q_1 and 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 between Q_1 and Q_2 is then defined by

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

where f_i is a μ-density of Q_i, $i = 1, 2$. Note that the variational distance is bounded by the Hellinger distance. For an explanation of why we prefer the Hellinger distance and for further technical details, we refer to Reiss [(1993), Section 1.3].

THEOREM 1.2. *Choose $a_n > 0$ such that $a_n \to 0$, $na_n \to \infty$ as $n \to \infty$. For $\vartheta \in \mathbb{R}$, set*

$$\beta_n := \beta_n(\vartheta) := \frac{\vartheta a_n^{1/2}}{\log^2(na_n)}.$$

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

$$\text{(C)} \qquad \left| \frac{f_{\beta_n}(x)}{h_{\beta_n}(x)} - 1 \right| \le C\left(1 - H_{\beta_n}(x)\right)^{\delta}, \qquad x \in \left[x_0(\beta_n), \omega\left(F_{\beta_n}\right)\right),$$

for some $C, \delta > 0$, where $x_0(\beta_n) \le t_n = \log(na_n)$. Then, uniformly for ϑ in compact subsets of $\mathbb{R}$,

$$H\left(\mathscr{L}_{\beta_n}(N_n), \mathscr{L}_{\beta_n}(N_n^*)\right) = O\left(\frac{1}{\log(na_n)} + \frac{1}{(na_n)^{\delta} a_n^{1/2}} \right).$$

REMARKS. It is just an exercise to show that the rate $O(\log^{-1}(na_n) + (na_n)^{-\delta} a_n^{-1/2})$ in the preceding result is sharp. The second error term $(na_n)^{-\delta} a_n^{-1/2}$ can be dropped in case of equality $f_{\beta_n}(x) = h_{\beta_n}(x)$, $x \in [x_0(\beta_n), \omega(F_{\beta_n}))$. Furthermore, the proof of Theorem 1.2 shows that

$$H(\mathscr{L}_0(N_n), \mathscr{L}_0(N_n^*)) = O\left(\frac{1}{(na_n)^{\delta} a_n^{1/2}}\right).$$

The bound $O(1/\log(na_n)) = O(1/t_n)$ in the preceding result entails that $\tau(n)$ becomes sufficient at a rather slow rate as the sample size n increases. The obvious advice by Theorem 1.1, simply to drop the information contained in the excesses for small up to moderate n, could therefore be taken only with a grain of salt; see also the remarks at the end of the paper.

PROOF OF THEOREM 1.2. From Corollary 1.2.4(iv) in Falk, Hüsler and Reiss (1994) we obtain, if $x_0(\beta_n) \le t_n = \log(na_n) < \omega(H_{\beta_n})$,

$$\begin{aligned} &H\big(\mathscr{L}_{\beta_n}(N_n), \mathscr{L}_{\beta_n}(N_n^*)\big) \\ &\quad \le (E_0(\tau_n))^{1/2} H(B_{\beta_n}, B_0) \\ &\quad = n^{1/2}\left\{\int_0^\infty \left[f_{\beta_n}^{1/2}(y + t_n)\exp\left(\frac{y}{2}\right) - \big(1 - F_{\beta_n}(t_n)\big)^{1/2}\right]^2 \exp(-y)\, dy\right\}^{1/2}. \end{aligned}$$

From now on we suppress the index n of β_n. Elementary computations yield the following fact.

FACT 2. For $y \in [x_0(\beta), \omega(H_\beta))$ and some constant $C_1 > 0$,

$$\left|\frac{1 - F_\beta(y)}{1 - H_\beta(y)} - 1\right| \le C_1\big(1 - H_\beta(y)\big)^{\delta}.$$

Fact 2 implies the following fact.

FACT 3. We have

$$\begin{aligned} &\int_{1/|\beta|^{1/2}}^\infty \left[f_\beta^{1/2}(y + t_n)\exp\left(\frac{y}{2}\right) - \big(1 - F_\beta(t_n)\big)^{1/2}\right]^2 \exp(-y)\, dy \\ &\qquad = O\big((na_n)^{-2}\big). \end{aligned}$$

From Fact 3 we deduce that

$$\begin{aligned} &H\big(\mathscr{L}_{\beta_n}(N_n), \mathscr{L}_{\beta_n}(N_n^*)\big)^2 \\ &\quad \le n\left\{\int_0^{1/|\beta|^{1/2}}\left[f_\beta^{1/2}(y + t_n)\exp\left(\frac{y}{2}\right) - \big(1 - F_\beta(t_n)\big)^{1/2}\right]^2 \exp(-y)\, dy\right\} \\ &\qquad + O\big((na_n)^{-2}\big), \end{aligned}$$

and from Fact 2 and condition (C) we obtain

$$n\int_0^{1/|\beta|^{1/2}}\left[f_\beta^{1/2}(y+t_n)\exp\left(\frac{y}{2}\right)-\left(1-F_\beta(t_n)\right)^{1/2}\right]^2\exp(-y)\,dy$$

$$=n\int_0^{1/|\beta|^{1/2}}\left[h_\beta^{1/2}(y+t_n)\exp\left(\frac{y}{2}\right)\left(1+O\left(\left(1-H_\beta(y+t_n)\right)\right)^{\delta}\right)\right.$$

$$\left.-\left(1-H_\beta(t_n)\right)^{1/2}\left(1+O\left((na_n)^{-\delta}\right)\right)\right]^2\exp(-y)\,dy$$

$$=n\int_0^{1/|\beta|^{1/2}}\left[\exp\left\{\frac{y}{2}-\frac{1+\beta}{2\beta}\log(1+\beta(y+t_n))\right\}\left(1+O\left((na_n)^{-\delta}\right)\right)\right.$$

$$\left.-\exp\left\{-\frac{1}{2\beta}\log(1+\beta t_n)\right\}\left(1+O\left((na_n)^{-\delta}\right)\right)\right]^2\exp(-y)\,dy$$

$$=a_n^{-1}\exp\left(\frac{\beta t_n^2}{2}\right)$$

$$\times\int_0^{1/|\beta|^{1/2}}\left[\exp\left\{\left(\frac{\beta}{4}\right)(y^2+2yt_n)\right.\right.$$

$$\left.+O\left(\beta(y+t_n)+\beta^2(y+t_n)^3\right)\right\}\left(1+O\left((na_n)^{-\delta}\right)\right)$$

$$\left.-\exp\left\{O\left(\beta^2t_n^3\right)\right\}\left(1+O\left((na_n)^{-\delta}\right)\right)\right]^2\exp(-y)\,dy;$$

recall that $H_\beta(t_n)=O(1/n)$. From the expansion $|\exp(x)-1|\le 3|x|$, for $|x|\le 1$, we deduce that the preceding term is of order

$$a_n^{-1}\int_0^{1/|\beta|^{1/2}}\left[O\left(|\beta|\left(y^2+2yt_n\right)+\beta(y+t_n)+\beta^2(y+t_n)^3\right)\right.$$

$$\left.+O\left((na_n)^{-\delta}\right)\right]^2\exp(-y)\,dy$$

$$=a_n^{-1}O\left(|\beta t_n|^2+(na_n)^{-2\delta}\right)=O\left(t_n^{-2}+a_n^{-1}(na_n)^{-2\delta}\right),$$

which implies the assertion. □

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

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

Note that L_β with $\beta>0$ is the standard Pareto distribution and L_{-1} is, for example, the uniform distribution on $[-1,0]$; L_β can be obtained from H_β by the identity $L_\beta(x) = H_{1/\beta}(\beta(x-1))$ if $\beta > 0$ and $L_\beta(x) = H_{1/\beta}(-\beta(x+1))$ if $\beta < 0$.

In the following we modify the empirical point process of the excesses pertaining to the exceedances $V_1, \dots, V_{\tau(n)}$ over the threshold t_n (greater than 1 if $\beta > 0$ and between -1 and 0 if $\beta < 0$). We consider instead the process

$$M_n := \sum_{i=1}^{\tau(n)} \varepsilon_{V_i/|t_n|},$$

where, by Lemma 0.1, $\tau(n)$ is $B(n, 1 - L_\beta(t_n))$-distributed and independent of $V_1/|t_n|, V_2/|t_n|, \dots,$ which are iid with common d.f.

$$B_\beta(x) = 1 - \frac{1 - L_\beta(y|t_n|)}{1 - L_\beta(t_n)} = L_\beta(x).$$

The reason for replacing N_n by M_n is the fact that in the case $\beta \neq 0$ the excess d.f. B_β is stable in the sense that it is again equal to L_β. The d.f. $B_\beta(y) = L_\beta(y)$ has density

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

Now fix $\beta_0 \neq 0$ and choose the threshold

$$t_n = t_{n,\beta_0} := L_{\beta_0}^{-1}\left(1 - (na_n)^{-1}\right) = \operatorname{sign}(\beta_0)(na_n)^{1/\beta_0}$$

and the alternatives $\beta_n = \beta_n(\vartheta)$ such that

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

where again $a_n > 0$, $n \in \mathbb{N}$, satisfies $a_n \to 0$, $na_n \to \infty$ as $n \to \infty$.

Notice that the alternatives β_n converge to $\beta_0 \neq 0$ at a slower rate than in our previous considerations, that is, $a_n^{1/2}/\log(na_n)$ compared to $a_n^{1/2}/\log^2(na_n)$. It is therefore more difficult to distinguish between hypothesis and alternatives in the subfamilies $\{L_\beta: \beta < 0\}$ and $\{L_\beta: \beta > 0\}$ than to decide between $\beta = 0$ and $\beta \neq 0$ in the general class $\{H_\beta: \beta \in \mathbb{R}\}$.

THEOREM 1.3 (LAN). *For $\beta_0 \neq 0$ we have, with the preceding choice of alternatives and the particular threshold $t_{n,\beta_0} = \operatorname{sign}(\beta_0)(na_n)^{1/\beta_0}$,*

$$\log\left\{\frac{d\mathscr{L}_{\beta_n}(M_n)}{d\mathscr{L}_{\beta_0}(M_n)}\right\}(M_n)$$
$$= \vartheta a_n^{1/2}\left(\tau(n) - a_n^{-1}\right) - \frac{\vartheta^2}{2} + o_{P_{\beta_0}}(1) \to_{\mathscr{D}_{\beta_0}} N\left(-\frac{\vartheta^2}{2}, \vartheta^2\right).$$

Recall that under β_0 the r.v. $\tau(n) = M_n(\mathbb{R})$ is $B(n, 1 - L_{\beta_0}(t_{n,\beta_0})) = B(n, 1/(na_n))$-distributed. An asymptotically optimal level α test for β_n against β_0 is, by Theorem 1.3, again given by φ_n as in (2) with N_n replaced by M_n, and asymptotic power function $\psi(\vartheta)$ defined in (3).

PROOF OF THEOREM 1.3. As in the proof of Theorem 1.1 we have, with $t_n = t_{n,\beta_0}$,

$$\log\left(\frac{d\mathscr{L}_{\beta_n}(M_n)}{d\mathscr{L}_{\beta_0}(M_n)}\right)(\mu) = \int \log\left\{\frac{b_{\beta_n}(y)}{b_{\beta_0}(y)}\right\}\mu(dy)$$

$$+ \mu(\mathbb{R})\log\left\{\frac{1-L_{\beta_n}(t_n)}{1-L_{\beta_0}(t_n)}\right\} + \left(n-\mu(\mathbb{R})\log\left(\frac{L_{\beta_n}(t_n)}{L_{\beta_0}(t_n)}\right)\right).$$

Elementary computations yield the following fact.

FACT 4. We have $L_{\beta_n}(t_n) - L_{\beta_0}(t_n) = -(na_n)^{-1}(\vartheta a_n^{1/2} + O(a_n))$.

With $\mu(\mathbb{R})$ replaced by $\tau(n)$, and $E_{\beta_0}(\tau(n)) = n(1-L_{\beta_0}(t_n)) = 1/a_n$ under β_0, we obtain, as in the proof of Theorem 1.1 under β_0,

$$\tau(n)\log\left\{\frac{1-L_{\beta_n}(t_n)}{1-L_{\beta_0}(t_n)}\right\} + (n-\tau(n))\log\left\{\frac{L_{\beta_n}(t_n)}{L_{\beta_0}(t_n)}\right\}$$

$$= \left(\tau(n) - a_n^{-1}\right)a_n^{1/2}\vartheta - \frac{\vartheta^2}{2} + o_{P_0}(1);$$

recall that $L_{\beta_0}(t_n) = 1-(na_n)^{-1}$ and that, by Fact 4, $L_{\beta_n}(t_n) - L_{\beta_0}(t_n) = \vartheta a_n^{1/2}/(na_n) + O(1/n)$. Finally, we have, under β_0,

$$\int \log\left\{\frac{b_{\beta_n}(y)}{b_{\beta_0}(y)}\right\}M_n(dy) = \sum_{i=1}^{\tau(n)} \log\left\{\left(\frac{\beta_n}{\beta_0}\right)\left(\left|\frac{V_i}{t_n}\right|\right)^{\beta_0-\beta_n}\right\}$$

$$=_{\mathscr{D}_{\beta_0}} \sum_{i=1}^{\tau(n)} \log\left\{\left(\frac{\beta_n}{\beta_0}\right)|W_i|^{\beta_0-\beta_n}\right\},$$

where $W_1, W_2, \ldots$ are iid with common d.f. L_{β_0} and independent of $\tau(n)$, which is $B(n, 1/(na_n))$-distributed. The last term equals

$$\sum_{i=1}^{\tau(n)}\left(\log\left\{1+\frac{\beta_n-\beta_0}{\beta_0}\right\} + (\beta_0-\beta_n)\log(|W_i|)\right)$$

$$= \tau(n)O\left(\frac{a_n}{\log(na_n)}\right) + (\beta_n-\beta_0)\sum_{i=1}^{\tau(n)}\left(\beta_0^{-1} - \log(|W_i|)\right).$$

Now observe that $E_{\beta_0}(\log(|W_i|)) = 1/\beta_0$ and $E_{\beta_0}(\log^2(|W_i|)) = 2/|\beta_0|^2$. The central limit theorem together with the fact that $\tau(n)a_n$ converges to 1 in probability implies that the final line above is of order $o_{P_{\beta_0}}(1)$, which completes the proof of Theorem 1.3. □

Theorem 1.3 remains true if we replace $\{L_\beta\colon \beta \neq 0\}$ by the δ-neighborhood

$$Q(C,\delta) = \left\{ \begin{array}{l} F_\beta\colon F_\beta \text{ is a d.f. with } \omega(F_\beta) = \omega(L_\beta) \text{ having a density } f_\beta \\ \text{on } \left[x_0(\beta), \omega(F_\beta)\right) \text{ for some } x_0(\beta) < \omega(F_\beta) \text{ such that} \\ \left| \dfrac{f_\beta(x)}{l_\beta(x)} - 1 \right| \le C\left(1 - L_\beta(x)\right)^\delta,\ x \in \left[x_0(\beta), \omega(F_\beta)\right) \end{array} \right\},$$

and require that $x_0(\beta_n) \le t_{n,\beta_0} = L_{\beta_0}^{-1}(1-(na_n)^{-1})$. Then again $(\tau(n) - E_{\beta_0}(\tau(n)))a_n^{1/2} = (\tau(n) - a_n^{-1})a_n^{1/2} + o_{P_{\beta_0}}(1)$ is the central sequence for testing β_0 against $\beta_n(\vartheta)$, with $(\beta_n - \beta_0)/\beta_0 = \vartheta a_n^{1/2}/\log(na_n)$, $n \in \mathbb{N}$, if in addition the sequence $a_n > 0$, $n \in \mathbb{N}$, satisfies $a_n^{1/2}(na_n)^\delta \to \infty$ as $n \to \infty$.

In the following we will establish the analogous result to Theorem 1.2 in the case of $\beta_0 \neq 0$. We will establish a bound for the Hellinger distance between $\mathscr{L}_{\beta_n}(M_n) = \mathscr{L}_{\beta_n}(\sum_{i=1}^{\tau(n)} \varepsilon_{V_i/|t_{n,\beta_0}|})$ and $\mathscr{L}_{\beta_n}(M_n^*) = \mathscr{L}_{\beta_n}(\sum_{i=1}^{\tau(n)} \varepsilon_{W_i})$, where F_{β_n} is in $Q(C,\delta)$, and M_n^* is obtained from M_n by replacing $V_1/|t_{n,\beta_0}|$, $V_2/|t_{n,\beta_0}|, \dots$ by $W_1, W_2, \dots$, which are iid with common d.f. L_{β_0} and independent of $\tau(n) = M_n(\mathbb{R})$. The rate for the sufficiency of $\tau(n)$, obtained in the following result, coincides with that in Theorem 1.2.

THEOREM 1.4. *Choose $a_n > 0$ such that $a_n \to 0$, $na_n \to \infty$ as $n \to \infty$, and choose for $\vartheta \in \mathbb{R}$ the sequence $\beta_n = \beta_n(\vartheta)$ such that*

$$\frac{\beta_0 - \beta_n}{\beta_0} = \frac{\vartheta a_n^{1/2}}{\log(na_n)}.$$

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

$$\text{(C)} \qquad \left| \frac{f_{\beta_n}(x)}{l_{\beta_n}(x)} - 1 \right| \le C\left(1 - L_{\beta_n}(x)\right)^\delta, \qquad x \in \left[x_0(\beta_n), \omega(F_{\beta_n})\right)$$

for some $C, \delta > 0$, where

$$x_0(\beta_n) \le t_{n,\beta_0} = L_{\beta_0}^{-1}\left(1 - (na_n)^{-1}\right) = \operatorname{sign}(\beta_0)(na_n)^{1/\beta_0}, \qquad n \in \mathbb{N}.$$

Then

$$H\left(\mathscr{L}_{\beta_n}(M_n), \mathscr{L}_{\beta_n}(M_n^*)\right) = O\left(\frac{1}{\log(na_n)} + \frac{1}{(na_n)^\delta a_n^{1/2}}\right).$$

REMARK. In case of equality $f_{\beta_n}(x) = l_{\beta_n}(x)$, the error term $(na_n)^{-\delta}a_n^{-1/2}$ can again be dropped. Equally, the proof of Theorem 1.4 shows that

$$H\left(\mathscr{L}_{\beta_0}(M_n), L_{\beta_0}(M_n^*)\right) = O\left(\frac{1}{(na_n)^\delta a_n^{1/2}}\right).$$

PROOF OF THEOREM 1.4. For the sake of a clear presentation, in the following we will write t_n in place of t_{n,β_0}. Repeating the arguments of the proof of Theorem 1.1, we obtain

$$H\big(\mathscr{L}_{\beta_n}(M_n), \mathscr{L}_{\beta_n}(M_n^*)\big) \le n^{1/2}\left\{\int_{\mathbb{R}}\left[|t_n|^{1/2}\frac{f_{\beta_n}^{1/2}(y|t_n|)}{l_{\beta_0}^{1/2}(y)} - \big(1-F_{\beta_n}(t_n)\big)^{1/2}\right]^2 l_{\beta_0}(y)\,dy\right\}^{1/2}$$

and the following fact.

FACT 5. For $y \in [x_0(\beta_n), \omega(F_{\beta_n}))$ and some constant $C_1 > 0$,

$$\left|\left(\frac{1-F_{\beta_n}(y)}{1-L_{\beta_n}(y)}\right)^{1/2} - 1\right| \le C_1\big(1-L_{\beta_n}(y)\big)^{\delta}.$$

Consequently, we have, by condition (C) and Fact 5,

$$\begin{aligned}
&n\int_{\mathbb{R}}\left[|t_n|^{1/2}\frac{f_{\beta_n}^{1/2}(y|t_n|)}{l_{\beta_0}^{1/2}(y)} - \big(1-F_{\beta_n}(t_n)\big)^{1/2}\right]^2 l_{\beta_0}(y)\,dy\\
&\quad= n\int_{\mathbb{R}}\Bigg[|t_n|^{1/2}\frac{l_{\beta_n}^{1/2}(y|t_n|)}{l_{\beta_0}^{1/2}(y)}\left\{1+\left(\frac{f_{\beta_n}^{1/2}(y|t_n|)}{l_{\beta_n}^{1/2}(y|t_n|)}-1\right)\right\}\\
&\qquad\qquad -\big(1-L_{\beta_n}(t_n)\big)^{1/2}\left\{1+\left(\left(\frac{1-F_{\beta_n}(t_n)}{1-L_{\beta_n}(t_n)}\right)^{1/2}-1\right)\right\}\Bigg]^2 l_{\beta_0}(y)\,dy\\
&\quad= n\int_{\mathbb{R}}\Bigg[|t_n|^{1/2}\frac{l_{\beta_n}^{1/2}(y|t_n|)}{l_{\beta_0}^{1/2}(y)} - \big(1-L_{\beta_n}(t_n)\big)^{1/2}\\
&\qquad\qquad +\left\{|t_n|^{1/2}\frac{l_{\beta_n}^{1/2}(y|t_n|)}{l_{\beta_0}^{1/2}(y)} + \big(1-L_{\beta_n}(t_n)\big)^{1/2}\right\}\\
&\qquad\qquad\qquad \times O\Big(\big(1-L_{\beta_n}(t_n)\big)^{\delta}\Big)\Bigg]^2 l_{\beta_0}(y)\,dy\\
&\quad= n|t_n|^{-\beta_n}\int_{\mathbb{R}}\Bigg[\left(\frac{\beta_n}{\beta_0}\right)^{1/2}|y|^{(\beta_0-\beta_n)/2} - 1\\
&\qquad\qquad +\left\{\left(\frac{\beta_n}{\beta_0}\right)^{1/2}|y|^{(\beta_0-\beta_n)/2}+1\right\}O\big(|t_n|^{-\delta\beta_n}\big)\Bigg]^2 l_{\beta_0}(y)\,dy.
\end{aligned}$$

The definition of β_n implies $|t_n|^{-\beta_n} = O((na_n)^{-1})$. By the expansion

$(1+\varepsilon)^{1/2} = 1 + O(\varepsilon)$, $\varepsilon \to 0$, the above integral equals therefore

$$O\big(a_n^{-1}\big)\int_{\mathbb{R}}\Bigg[\bigg(1+\frac{\beta_n-\beta_0}{\beta_0}\bigg)^{1/2}|y|^{(\beta_0-\beta_n)/2}-1$$

$$+\bigg\{\bigg(\frac{\beta_n}{\beta_0}\bigg)^{1/2}|y|^{(\beta_0-\beta_n)/2}+1\bigg\}O\Big((na_n)^{-\delta}\Big)\Bigg]^2 l_{\beta_0}(y)\,dy$$

$$= O\big(a_n^{-1}\big)\int_{\mathbb{R}}\Big[(1+O(\beta_n-\beta_0))|y|^{(\beta_0-\beta_n)/2}-1$$

$$+O\Big((na_n)^{-\delta}\big(1+|y|^{(\beta_0-\beta_n)/2}\big)\Big)\Big]^2 l_{\beta_0}(y)\,dy$$

$$= O\big(a_n^{-1}\big)\int_{\mathbb{R}}\Bigg[|y|^{(\beta_0-\beta_n)/2}-1+O\Big((na_n)^{-\delta}\Big)$$

$$+O\Bigg(\bigg\{\frac{a_n^{1/2}}{\log(na_n)}+\frac{1}{(na_n)^{\delta}}\bigg\}|y|^{(\beta_0-\beta_n)/2}\Bigg)\Bigg]^2 l_{\beta_0}(y)\,dy.$$

Now observe that

$$\int_{\mathbb{R}}\Big[|y|^{(\beta_0-\beta_n)/2}-1\Big]^2 l_{\beta_0}(y)\,dy=\frac{(\beta_0-\beta_n)^2}{\beta_n(\beta_0+\beta_n)}=O\bigg(\frac{a_n}{\log(na_n)^2}\bigg)$$

and that $\int_{\mathbb{R}}|y|^{\beta_0-\beta_n}l_{\beta_0}(y)\,dy = O(1)$ if n is large, since $\beta_0-\beta_n\to 0$ as $n\to\infty$. The above integral is therefore of order $O(a_n/\log(na_n)^2+(na_n)^{-2\delta})$, which implies the assertion. □

2. Adding a scale parameter. In the following we extend the statistical models of Section 1 by adding a scale parameter $c>0$; that is, we consider a sequence $Y_1, Y_2, \dots$ of independent r.v.'s with common d.f. F whose upper tail belongs to a parametric family

$$1-F(x)=1-F_\beta(cx),\qquad x\ge x_0=x_0(F),$$

with the scale and the shape parameter $c>0$ and $\beta\in\Theta\subset\mathbb{R}$ as well as the root x_0 being unknown.

The results of this section parallel those of the previous section, as it turns out that again only the number $\tau(n)$ of excesses carries asymptotically all the information contained in N_n. Thus, even in case of the more detailed testing problem (β_0, c_0) against (β_n, c_n), the exceedances $V_1,\dots,V_{\tau(n)}$ over t_n do not contribute asymptotically to its solution. As a consequence, statistical testing can be based in our model on $\tau(n)$ alone without any loss of asymptotic efficiency.

In the following we specify our statistical model. Again denote by H_β the GPD-d.f. in its von Mises parametrization and define the scale-shifted version, for $c > 0$, by

$$H_{\beta,c}(x) := H_\beta(cx) = 1 - (1+\beta cx)^{-1/\beta}, \qquad \begin{cases} x \ge 0, & \text{if } \beta \ge 0, \\ 0 \le x \le -1/(c\beta), & \text{if } \beta < 0, \end{cases}$$

Again interpret $H_{0,c}(x)$ as $H_{0,c}(x) = \lim_{\beta\to 0} H_{\beta,c}(x) = 1 - \exp(-cx)$, $x \ge 0$. Let $h_{\beta,c}$ denote the density of $H_{\beta,c}$.

The d.f. of the excess distribution over the threshold t_n now becomes

$$\begin{aligned} B_{\beta,c}(x) & \\ &:= 1 - \frac{1 - H_{\beta,c}(x+t_n)}{1 - H_{\beta,c}(t_n)} \\ &= 1 - \left(1 + \frac{\beta cx}{1+\beta ct_n}\right)^{-1/\beta} \quad \text{for } \begin{cases} x \ge 0, & \text{if } \beta \ge 0, \\ 0 \le x \le -(\beta c)^{-1}(1+\beta ct_n), & \text{if } \beta < 0, \end{cases} \end{aligned}$$

provided t_n satisfies $0 < H_{\beta,c}(t_n) = H_\beta(ct_n) < 1$, which we assume in the following. Observe that $B_{0,c}(x) = H_{0,c}(x) = H_0(cx) = 1 - \exp(-cx)$, $x \ge 0$. The excess d.f. $B_{\beta,c}$ has density

$$b_{\beta,c}(x) = \frac{c}{1+\beta ct_n}\left(1 + \frac{\beta cx}{1+\beta ct_n}\right)^{-1/\beta - 1}$$
$$\text{for } \begin{cases} x \ge 0, & \text{if } \beta \ge 0, \\ 0 \le x \le -(\beta c)^{-1}(1+\beta ct_n), & \text{if } \beta < 0, \end{cases}$$

with $b_{0,c}(x) = ch_0(cx) = c\exp(-cx)$, $x \ge 0$.

In the following we consider at first the hypothesis.

$$(\beta_0, c_0) = (0, 1);$$

then we will consider a general $\beta_0 \ne 0$, but for the sake of simplicity we will always keep $c_0 = 1$. This reduction can obviously be achieved for a general hypothetical value $c_0 > 0$ by simply multiplying the initial observations $Y_1, Y_2, \dots$ by c_0 and considering $c_0 Y_1, \dots, c_0 Y_n$ instead. We suppose implicitly that this data manipulation has already been carried out.

As a consequence, the alternatives c_n which we will consider will always approach 1 as n increases. To be precise, choose the threshold $t_n := \log(na_n)$ and the alternatives

$$\beta_n := \beta_n(\vartheta) := 2\vartheta a_n^{1/2}/t_n^2, \qquad c_n := c_n(\xi) := 1 - \xi a_n^{1/2}/t_n,$$

for $\vartheta, \xi \in \mathbb{R}$, where the sequence $a_n > 0$, $n \in \mathbb{N}$, again satisfies $a_n \to 0$, $na_n \to \infty$ as $n \to \infty$. Note that the definitions of t_n and β_n coincide with those of Theorem 1.1.

Suppose now that we are given a family $\{F_{\beta,c}: \beta \in \mathbb{R},\ c > 0\}$ of d.f.'s such that $F_{0,1}(x) = H_{0,1}(x) = H_0(x) = 1 - \exp(-x)$ for $x \ge x_0 > 0$ and $F_{\beta_n,c_n}(x) = H_{\beta_n,c_n}(x)$ for $x \ge t_n$ if n is large. Then we have the following result which parallels Theorem 1.1. Its proof is completely analogous to that of Theorem

1.1 by utilizing the equation $H_{\beta_n, c_n}(t_n) - H_{0,1}(t_n) = -(na_n)^{-1}((\vartheta + \xi)a_n^{1/2} + o(1))$ in place of Fact 1.

THEOREM 2.1 (LAN). *Under the preceding model we have, under $F_{0,1}$, the expansion*

$$\log\left\{\frac{d\mathscr{L}_{\beta_n, c_n}(N_n)}{d\mathscr{L}_{0,1}(N_n)}\right\}(N_n) = (\vartheta + \xi)a_n^{1/2}\left(\tau(n) - a_n^{-1}\right) - \frac{(\vartheta + \xi)^2}{2} + o_{P_{0,1}}(1)$$

$$\to_{\mathscr{D}_{0,1}} N\left(-\frac{(\vartheta + \xi)^2}{2}, (\vartheta + \xi)^2\right).$$

Note that $\tau(n) = N_n((t_n, \infty))$ is $B(n, 1 - F_{0,1}(t_n)) = B(n, 1/(na_n))$-distributed under $F_{0,1}$ if n is large. This immediately implies the asymptotic normality in the above result.

The preceding result shows that we can distinguish between $(\beta_n, c_n) = (\beta_n(\vartheta), c_n(\xi))$ and $(0,1)$ asymptotically if and only if $\vartheta \neq -\xi$; as in the case $\vartheta = -\xi$, the limiting distributions of the log-likelihood ratio

$$\log\left\{d\mathscr{L}_{\beta_n, c_n}(N_n)/d\mathscr{L}_{0,1}(N_n)\right\}(N_n)$$

under (β_n, c_n) and under $(0,1)$ coincide. Asymptotically optimal tests for testing $(\beta_n(\vartheta), c_n(\xi))$ against $(0,1)$, where $\vartheta + \xi \neq 0$, are again given by (2) with sign(ϑ) replaced by sign$(\vartheta + \xi)$ and asymptotic power function $\psi(\vartheta + \xi)$ as in (3).

Theorem 2.1 remains true, if we replace $H_{\beta_n, c_n}(x)$, $x \geq t_n$, $H_{0,1}(x)$, $x \geq x_0$, by δ-neighborhoods F_{β_n, c_n} and $F_{0,1}$, having densities f_{β_n, c_n} and $f_{0,1}$ on $[t_n, \infty)$ and $[x_0, \infty)$ such that

$$\left|\frac{f_{\beta_n, c_n}(x)}{h_{\beta_n, c_n}(x)} - 1\right| \leq C\left(1 - H_{\beta_n, c_n}(x)\right)^{\delta}, \qquad x \geq t_n,$$

and

$$\left|\frac{f_{0,1}(x)}{h_{0,1}(x)} - 1\right| \leq C\left(1 - H_{0,1}(x)\right)^{\delta}, \qquad x \geq x_0,$$

provided that the sequence a_n satisfies in addition $a_n^{1/2}(na_n)^{\delta} \to \infty$ as $n \to \infty$.

The following result parallels Theorem 1.2. It provides a bound for the information contained in $\tau(n)$ about both parameters β_n and c_n. We compare again the distribution of the point process $N_n = \sum_{i=1}^{\tau(n)} \varepsilon_{V_i - t_n}$ under F_{β_n, c_n} with that of the point process $N_n^* = \sum_{i=1}^{\tau(n)} \varepsilon_{W_i}$, where $\tau(n) = N_n((0, \infty))$ is kept, but $W_1, W_2, \ldots$ are independent and standard exponential distributed r.v.'s. Its proof is completely analogous to that of Theorem 1.2 and therefore omitted.

THEOREM 2.2. *Choose $a_n > 0$ such that $a_n \to 0$, $na_n \to \infty$ as $n \to \infty$. For ϑ, $\xi \in \mathbb{R}$, set*

$$\beta_n = \beta_n(\vartheta) = \vartheta a_n^{1/2}/t_n^2, \qquad c_n = c_n(\xi) = 1 - \xi a_n^{1/2}/t_n,$$

where $t_n = \log(na_n)$. *Suppose that* $\omega(F_{\beta_n, c_n}) = \omega(H_{\beta_n, c_n})$ *and that*

$$\text{(C)} \qquad \left| \frac{f_{\beta_n, c_n}(y)}{h_{\beta_n, c_n}(y)} - 1 \right| \le C\big(1 - H_{\beta_n, c_n}(y)\big)^{\delta}, \qquad y \in \big[t_n, \omega(H_{\beta_n, c_n})\big),$$

for some $C, \delta > 0$. *Then*

$$H\big(\mathscr{L}_{\beta_n, c_n}(N_n), \mathscr{L}_{\beta_n, c_n}(N_n^*)\big) = O\left(\frac{1}{\log(na_n)} + \frac{1}{(na_n)^{\delta} a_n^{1/2}} \right).$$

REMARK. The error term $(na_n)^{-\delta} a_n^{1/2}$ can again be dropped if $f_{\beta_n, c_n}(y) = h_{\beta_n, c_n}(y)$, $y \in [t_n, \omega(H_{\beta_n, c_n}))$. On the other hand we have

$$H(\mathscr{L}_{0,1}(N_n), \mathscr{L}_{0,1}(N_n^*)) = O\left(\frac{1}{(na_n)^{\delta} a_n^{1/2}} \right).$$

Consider next the case $\beta_0 \ne 0$ with underlying tail distribution in the form

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

with scale parameter $c > 0$. However, in this case the POT-method cancels the scale parameter c, when we consider again the process

$$M_n = \sum_{i=1}^{\tau(n)} \varepsilon_{V_i / |t_n|}$$

pertaining to the exceedances $V_1, \dots, V_{\tau(n)}$ over the threshold t_n. This is immediate from

$$P_{\beta, c}\left\{ \frac{V_i}{|t_n|} \le x \right\} = 1 - \frac{1 - L_{\beta, c}(x|t_n|)}{1 - L_{\beta, c}(t_n)} = L_{\beta}(x), \qquad x \in \mathbb{R},$$

provided $0 < L_{\beta}(ct_n) < 1$.

As a consequence, the excesses cannot contribute any information about the underlying scale parameter within this approach. It is therefore clear that $\tau(n)$ plays an even more predominant role, as it not only carries asymptotically all the information from N_n about the shape parameter β, but it contains the complete information about the scale parameter $c > 0$ for finite sample size n as well. In order not to overload this paper with too many technicalities, however, we drop further details.

One clearly wonders about the information which the excesses themselves contribute to the knowledge about the underlying parameters, as the preceding results show that in those models their number $\tau(n)$ is already asymptotically sufficient. One way to upgrade the excesses is to consider the two-parameter problem (β, c) as before, but with the scale shift c being regarded as a nuisance parameter. The popular Hill estimator of the extreme value index β [Hill (1975)], which has been extensively studied in the literature [cf.

Falk, Hüsler and Reiss (1994), Section 2.4, and the references cited therein], is, for example, scale invariant. If one now investigates the testing problem $(\beta_0, 1)$ against (β_n, c_n), where c_n is such that (β_n, c_n) is some least favorable alternative to $(\beta_0, 1)$, then the excesses carry asymptotically the complete information about the underlying shape parameters. However, this is work still in progress and will be published in a subsequent paper.

Acknowledgments. The author is grateful to Frank Marohn for helpful discussions and to the referees for their careful reading of the manuscript and their detailed comments, from which the paper has benefited a lot.

REFERENCES

BALKEMA, A. A. and DE HAAN, L. (1974). Residual lifetime 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. In *Extreme Value Theory. Lecture Notes in Statist.* **51** 181–190. Springer, New York.

DALEY, D. J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, New York.

DAVISON, A. C. and SMITH, R. L. (1990). Models for exceedances over high thresholds (with discussion). *J. Roy. Statist. Soc. Ser. B* **52** 393–442.

DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes. Math. Centre Tracts* **32**. Mathematisch Centrum, Amsterdam.

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., 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, FL.

GOMES, M. I. (1989). Comparison of extremal models through statistical choice in multidimensional backgrounds. *Extreme Value Theory. Lecture Notes in Statist.* **51** 191–203. Springer, New York.

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.

HÖPFNER, R. (1994). A note on parameter estimation in stable processes. Unpublished manuscript. (Univ. Paris VI, preprint.)

HÖPFNER, R. and JACOD, J. (1993). Some remarks on the joint estimation of the index and the scale parameter for stable processes. Unpublished manuscript. (Univ. Paris VI, preprint.)

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.

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

LE CAM, L. (1986). *Asymptotic Methods in Statistical Decision Theory.* Springer, New York.

LE CAM, L. and YANG, G. L. (1990). *Asymptotics in Statistics* (*Some Basic Concepts*). 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. (1994a). On testing the exponential and Gumbel distribution. In *Extreme Value Theory and Applications* (J. Galambos, J. Lechner and E. Simio, eds.) 159–174. Kluwer, Dordrecht.

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

REISS, R.-D. (1989). *Approximate Distributions of Order Statistics* (*with Applications to Nonparametric Statistics*). Springer, New York.

REISS, R.-D. (1993). *A Course on Point Processes*. Springer, New York.

RYCHLIK, T. (1992). Weak limit theorems for stochastically largest order statistics. In *Order Statistics and Nonparametrics: Theory and Applications.* (P. K. Sen and I. A. Salama, eds.) 141–154. North-Holland, Amsterdam.

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

STRASSER, H. (1985). *Mathematical Theory of Statistics.* de Gruyter, Berlin.

TEUGELS, J. L. (1984). Extreme values in insurance mathematics. In *Statistical Extremes and Applications* (J. Tiago de Oliveira, ed.) 253–259. Reidel, Dordrecht.

WEI, X. (1992). Asymptotically efficient estimation of the index of regular variation. Ph.D. dissertation, Univ. Michigan.

MATHEMATISCH–GEOGRAPHISCHE FAKULTÄT
KATHOLISCHE UNIVERSITÄT EICHSTÄTT
85071 EICHSTÄTT
GERMANY