The Annals of Statistics
1984, Vol. 12, No. 4, 1467–1487

TAIL ESTIMATES MOTIVATED BY EXTREME VALUE THEORY

BY RICHARD DAVIS[1] AND SIDNEY RESNICK[2]

Colorado State University

An estimate of the upper tail of a distribution function which is based on the upper m order statistics from a sample of size n ($m \to \infty$, $m/n \to 0$ as $n \to \infty$) is shown to be consistent for a wide class of distribution functions. The empirical mean residual life of the log transformed data and the sample $1 - m/n$ quantile play a key role in the estimate. The joint asymptotic behavior of the empirical mean residual life and sample $1 - m/n$ quantile is determined and rates of convergence of the estimate to the tail are derived.

1. Introduction. This paper deals with the problem of estimating the tail of a distribution function (df) $\bar{F}(x) := 1 - F(x)$ for large x based on a random sample. Many of the proposed estimators of the tail (cf. Hill, 1975 and Breiman, Stone and Ginns, 1979) assume that $\bar{F}(x)$ belongs to a given parametric family (typically Pareto or exponential) for all x greater than some predetermined value x_0. The parameters are then estimated by maximizing the likelihood based on the observations which exceed x_0 (see DuMouchel, 1983, and the references therein). Two obvious problems in this approach are the choice of a parametric family and the threshold value x_0. We propose an estimate of $\bar{F}$ without appealing to the likelihood principle which is applicable to a wide class of distribution functions. The proposed estimator is motivated by ideas from classical extreme value theory and in some instances coincides with estimators given by other authors. Extreme value theory also is a clear influence on the estimator of Pickands (1975).

A df F belongs to the domain of attraction of an extreme value distribution G, if there exist constants $a_n > 0$, b_n such that $n\bar{F}(a_n x + b_n) \to -\log G(x)$ for all x with $G(x) > 0$. We shall assume that $-\log G(x) = (1 + x/\alpha)^{-\alpha}$ where $0 < \alpha \le \infty$, the case $\alpha = \infty$ corresponding to $-\log G(x) = e^{-x}$. The upper tail of F is then exponential or Pareto-like depending on whether α is infinite or finite. Under the assumptions given in Section 2, we show that

$$\sup_{x \ge b(t)} | t\bar{F}(x) - (x/b(t))^{-1/a^*(t)} | \to 0 \quad \text{as} \quad t \to \infty \tag{1.1}$$

where $b(t)$ is the $1 - t^{-1}$ quantile of F and $a^*(t) = t \int_{\log b(t)}^{\infty} \bar{F}(e^t)dt$ is the mean residual life of the df $F(e^t)$ evaluated at $\log b(t)$. In particular $a^*(t) \to \alpha^{-1}$ as $t \to \infty$ for all $0 < \alpha \le \infty$ (see Section 2). From a random sample $X_1, \cdots, X_n$, the estimates of the parameters $b(t)$ and $a^*(t)$ are based on the upper $m = m(n)$

Received February 1984; revised June 1984.

[1] Research supported by NSF Grant MCS08202335 and AFOSR F49620 82 C 0009.

[2] Research supported by NSF Grant MCS08202335.

AMS 1980 *subject classifications.* Primary 62G05; secondary 62G30, 62F12.

Key words and phrases. Tail estimation, regular variation, Pareto distributions, mean residual life, intermediate order statistics.

order statistics where m is a sequence of integers chosen such that $m \to \infty$ and $m/n \to 0$. This idea is similar to that of Weissman (1978) (see also Boos, 1984, for empirical results on the Weissman article) except that he assumes m is held fixed. Replacing t by n/m in (1.1), the parameter values $b(n/m)$ and $a^*(n/m)$ are naively estimated by their empirical counterparts. That is if $X_{(1)} > X_{(2)} > \cdots > X_{(n)}$ denotes the decreasing order statistics, then $b(n/m)$ is estimated by the empirical $1 - m/n$ quantile, $X_{(m+1)}$, and $a^*(n/m)$ is estimated by the empirical mean residual life of the log-transformed data evaluated at $\log X_{(m+1)}$ which is equal to

$$\widehat{a^*}(n/m) = m^{-1} \textstyle\sum_{i=1}^{m} (\log (X_{(i)}) - \log(X_{(m+1)})). \tag{1.2}$$

This leads to a Pareto tail estimate of $\bar{F}(x)$ given by

$$(m/n)(x/X_{(m+1)})^{-1/\widehat{a^*}(n/m)}. \tag{1.3}$$

This estimator of the tail is essentially the same as the one proposed in Hill (1975) (see also Breiman et al., 1979) where it is assumed that $\bar{F}(x) = cx^{-\alpha}$ for $x > x_0$ with x_0 known.

In Sections 4 and 5 the asymptotic properties of the estimators of $b(n/m)$ and $a^*(n/m)$ are derived. In particular it is shown that

$$\sqrt{m}\left(\frac{\widehat{a^*}(n/m)}{a^*(n/m)} - 1, \quad \frac{\log X_{(m+1)} - \log b(n/m)}{a^*(n/m)}\right)$$

converges jointly in distribution to two independent normal (0, 1) random variables. Consistency and rates of convergence of (1.3) to $\bar{F}$ are derived in Section 3.

2. Preliminaries. We now discuss our assumptions in more detail. We will always assume that $F(x)$, the distribution whose tail is to be estimated, satisfies $F(x) < 1$ for all x and that F has at least one derivative F' which serves as a density. Then $\bar{F}(x) = 1 - F(x)$ has the representation for $x > 0$

$$\bar{F}(x) = c \exp\left\{- \int_0^x (1/f_0(u))\, du\right\}, \quad c > 0 \tag{2.1}$$

where

$$f_0(u) = \bar{F}(u)/F'(u), \quad u > 0. \tag{2.2}$$

In order to guarantee that F is in the domain of attraction of an appropriate extreme value distribution we assume the validity of sufficient conditions due essentially to Von Mises (cf. de Haan, 1970, Section 2.7). For $\alpha < \infty$ assume

$$\lim_{x\to\infty} x^{-1} f_0(x) = \alpha^{-1}. \tag{2.3}$$

In this case set

$$b(t) = F^{\leftarrow}(1 - 1/t), \quad a_0(t) = f_0(b(t)) \tag{2.4}$$

where $F^{\leftarrow}(u)$ is the left continuous inverse of F and if $\alpha > 1$, we further define

$$a(t) = t \int_{b(t)}^{\infty} \bar{F}(s)\, ds$$

so that $a(t)$ is the mean residual life evaluated at $b(t)$. We then have as $t \to \infty$ for all $\alpha < \infty$,

$$\begin{aligned} t\bar{F}(a_0(t)x + b(t)) &= \exp\left(-\int_{b(t)}^{a_0(t)x+b(t)} \left(\frac{1}{f_0(u)}\right) du\right) \\ &= \exp\left(-\int_{1}^{(f_0(b(t))/b(t))x+1} \left(\frac{b(t)u}{f_0(b(t)u)}\right) u^{-1}\, du\right) \\ &\to \exp\left(-\int_{1}^{\alpha^{-1}x+1} \alpha u^{-1}\, du\right) \\ &= (1 + \alpha^{-1}x)^{-\alpha}, \end{aligned}$$

and provided $\alpha > 1$

$$\frac{a(t)}{a_0(t)} = \left(\frac{b(t)F'(b(t))}{\bar{F}(b(t))}\right)\left(\frac{\int_{b(t)}^{\infty} \bar{F}(s)\, ds}{b(t)\bar{F}(b(t))}\right) \to \alpha/(\alpha - 1) \tag{2.5}$$

(cf. de Haan, 1970, page 15).

For attraction to $\exp\{-e^{-x}\}$, referred to as the $\alpha = \infty$ case, we assume $\bar{F}$ has a representation for $x > 0$

$$\bar{F}(x) = c \exp\left\{-\int_{0}^{x} (h(u)/f_1(u))\, du\right\}, \quad c > 0 \tag{2.6}$$

where

$$\lim_{u\to\infty} h(u) = 1$$

and f_1 is differentiable with derivative f_1', satisfying

$$\lim_{u\to\infty} f_1'(u) = 0$$

(cf. de Haan, 1970, pages 92, 111). The condition (2.6) is implied by the usual Von Mises condition, $f_0'(u) \to 0$ as $u \to \infty$ in which case the most convenient choice of h and f_1 is $h \equiv 1$ and $f_1 = f_0$. The condition (2.6) also holds if (de Haan, 1970, page 110)

$$h(x) := F'(x) \int_{x}^{\infty} \bar{F}(s)\, ds/(\bar{F}(x))^2 \to 1 \quad \text{as } x \to \infty, \tag{2.7}$$

in which case we may take

$$f_1(x) = \int_{x}^{\infty} \bar{F}(s)\, ds/\bar{F}(x).$$

(In fact (2.6) is equivalent to (2.7).)

If (2.6) holds we may take

$$b(t) = F^{\leftarrow}(1 - 1/t), \quad a_1(t) = f_1(b(t)), \quad a(t) = t\int_{b(t)}^{\infty} \bar{F}(s)\,ds \tag{2.8}$$

and then since $\lim_{t\to\infty} b(t)^{-1}a_1(t) = \lim_{t\to\infty} b(t)^{-1}f_1(b(t)) = \lim_{t\to\infty} f_1'(b(t)) = 0$, $a_1(t)x + b(t) \to \infty$ and $\lim_{t\to\infty} f_1(a_1(t)x + b(t))/f_1(b(t)) = 1$, we get

$$\begin{aligned} t\bar{F}(a_1(t)x + b(t)) &= \exp\left\{-\int_{b(t)}^{a_1(t)x+b(t)} (h(u)/f_1(u))\,du\right\} \\ &= \exp\left\{-\int_0^x h(a_1(t)u + b(t))(f_1(b(t))/f_1(a_1(t)u + b(t)))\,du\right\} \\ &\to \exp\left\{-\int_0^x 1\,du\right\} = e^{-x}. \end{aligned}$$

Furthermore as $t \to \infty$ since $a_0(t) = f_0(b(t)) = f_1(b(t))/h(b(t))$

$$a(t) \sim a_1(t) \sim a_0(t) \tag{2.9}$$

(cf. de Haan, 1970, pages 84, 88).

When either (2.3) or (2.6) hold, we will say F is in an α-domain of attraction with the understanding that $\alpha = \infty$ refers to (2.6) being valid.

Finally, we note for later use that because $b(t) = F^{\leftarrow}(1 - 1/t)$ we have

$$b'(t) = \frac{\bar{F}(b(t))}{F'(b(t))t} = \frac{f_0(b(t))}{t} = \frac{a_0(t)}{t}. \tag{2.10}$$

Although we are being less than completely general in supposing (2.1) and either (2.3) or (2.6), we feel this is the appropriate level of generality for the tail estimation problem. The domains of attraction are not severely restricted by our assumptions (cf. de Haan, 1970, Section 2.7) and most common densities (normal, exponential, Cauchy, gamma, Pareto, $\cdots$) satisfy our conditions. Furthermore many of our results may be extended to the general case. This involves replacing the constant $c > 0$ in (2.1), (2.5) by a function $c(x)$ satisfying $\lim_{x\to\infty} c(x) = c > 0$ (de Haan, 1970; Balkema and de Haan, 1972).

In the tail estimation problem, it is advantageous for a number of reasons to be explained later to transform the original data by taking logarithms. (Negative observations, not relevant to estimating the right tail anyway, are neglected). If $X_1, \cdots, X_n$ is the original random sample we consider $X_1^*, \cdots, X_n^*$ where $X_i^* = \log X_i$. Then for $x > 0$, $\bar{F}^*(x) = P[X_1^* > x] = \bar{F}(e^x)$. If F is in a domain of attraction for $0 < \alpha \le \infty$, then F^* is in a domain of attraction with $\alpha = \infty$. In fact

$$f_0^*(x) = \frac{1 - F^*(x)}{(F^*)'(x)} = \frac{1 - F(e^x)}{e^x F'(e^x)} = \frac{f_0(e^x)}{e^x} \tag{2.11}$$

and

$$b^*(t) = \log b(t). \tag{2.12}$$

In the $\alpha < \infty$ case, if we set $h^*(x) = \alpha^{-1}/(f_0(e^x)e^{-x})$ and $f_1^* \equiv \alpha^{-1}$ we get for $x > 0$

$$\bar{F}^*(x) = c^*\exp\left\{ - \int_0^x \left(\frac{h^*(u)}{f_1^*(u)}\right) du\right\}$$

so that $\bar{F}^*$ is of form (2.6) and further from (2.9)

$$a^*(t) = t \int_{b^*(t)}^{\infty} \bar{F}^*(s)\, ds \sim a_1^*(t) = \alpha^{-1}. \tag{2.13}$$

If F is in a domain of attraction with $\alpha = \infty$, then (2.6) gives

$$\begin{aligned}\bar{F}^*(x) = \bar{F}(e^x) &= c \exp\left\{- \int_0^{e^x} \left(\frac{h(u)}{f_1(u)}\right) du\right\}\\ &= c^*\exp\left\{- \int_0^x \left(\frac{h(e^u)}{f_1(e^u)e^{-u}}\right) du\right\}\\ &= c^*\exp\left\{- \int_0^x \left(\frac{h^*(u)}{f_1^*(u)}\right) du\right\}\end{aligned}$$

where $h^*(u) = h(e^u)$, $f_1^*(u) = f_1(e^u)/e^u$. It is easy to check $h^* \to 1$ and $(f_1^*)' \to 0$ so that F^* satisfies (2.6). Furthermore in this case

$$a_1^*(t) = f_1^*(b^*(t)) = f_1(b(t))/b(t) = a_1(t)/b(t) \to 0 \tag{2.14}$$

as $t \to \infty$ and

$$a_1^*(t) \sim a^*(t) = t \int_{b^*(t)}^{\infty} \bar{F}^*(s)\, ds. \tag{2.15}$$

3. Approximations by Pareto tails. In this section we show that if F is in a domain of attraction for $0 < \alpha \le \infty$ as described in Section 2, then the tail $\bar{F}$ may be approximated by Pareto tails. By means of these approximations we will be able to show that our tail estimators are consistent. We begin with some notation. For a positive real number $g > 0$, define the distribution functions

$$F(g, x) = \begin{cases} 0 & \text{if } x < -g^{-1} \\ 1 - (1 + gx)^{-g^{-1}} & \text{if } x \ge -g^{-1} \end{cases}$$

and

$$F(-g, x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - (1 - gx)^{g^{-1}} & \text{if } 0 \le x < g^{-1} \\ 1 & \text{if } x \ge g^{-1}. \end{cases}$$

The distributions $F(\pm\, g, x)$ are close to a unit exponential distribution.

PROPOSITION 3.1. *For* $0 < g < 1$,

$$\sup_{x\ge 0} |\bar{F}(\pm\, g, x) - e^{-x}| \le (2 + g)e^{-2}g.$$

PROOF. The bound for $\sup_{x\ge 0} |\bar{F}(-g, x) - e^{-x}|$ is the one given by Hall and

Wellner (1979) and their method works almost without change to give the bound for $\sup_{x\geq 0}|F(g, x) - e^{-x}|$. □

Other approximations are collected in the next result.

Proposition 3.2. *For* $0 < g < 1$,

(i) $$\sup_{x\geq 0}|e^{-x(1-g)} - e^{-x}| \leq (g/(1-g))(1-g)^{1/g} \leq e^{-1}(g/(1-g)),$$

and

(ii) $$\sup_{x\geq 0}|e^{-x(1+g)} - e^{-x}| \leq (g/(1+g))(1+g)^{-1/g}.$$

Also for $0 < \alpha < \beta$

(iii) $$\sup_{x\geq 1}|x^{-\alpha} - x^{-\beta}| \leq \frac{|\beta - \alpha|}{\beta \wedge \alpha}(\alpha/\beta)^{\beta/(\beta-\alpha)} \leq e^{-1}|\beta - \alpha|/\beta \wedge \alpha.$$

Proof. (i) Set $\chi(x) = e^{-(1-g)x} - e^{-x} \geq 0$. It is easy to check that $\chi'(x)$ is positive for x near 0 and negative for large x. Thus it is evident that χ has a unique maximum and it occurs at the solution of $\chi'(x) = 0$ which is $-g^{-1}\log(1-g)$. The value of the maximum is thus

$$\chi(-g^{-1}\log(1-g)) = (g/(1-g))(1-g)^{1/g}.$$

The bounds for (ii) and (iii) can be obtained from (i). □

We now show how in the $\alpha = \infty$ case, the tail of a distribution in the domain of attraction of the double exponential $\exp\{-e^{-x}\}$, $x \in \mathbb{R}$, can be approximated by Pareto tails. This approximation is similar to one employed by A. A. Balkema, L. de Haan and S. Resnick in an unpublished manuscript. Begin by assuming F satisfies (2.6) and suppose $g(x) \geq 0$ is nonincreasing, $\lim_{x\to\infty} g(x) = 0$ and

(3.1) $$g(x) \geq |h(x) - 1| \vee |f_1'(x)|.$$

Such a function g always exists since we may take for instance

$$g(x) = \sup_{t\geq x}|h(t) - 1| \vee |f_1'(t)|.$$

Write for $t \geq b(u)$

$$f_1(t) - f_1(b(u)) = \int_{b(u)}^t f_1'(s)\,ds \leq \int_{b(u)}^t g(s)\,ds \leq g(b(u))(t - b(u))$$

$$\geq -\int_{b(u)}^t g(s)\,ds \geq -g(b(u))(t - b(u))$$

and similarly

$$1 - g(b(u)) \leq h(t) \leq 1 + g(b(u)).$$

Therefore for $t \geq b(u)$

$$\frac{1 - g(b(u))}{f_1(b(u)) + g(b(u))(t - b(u))} \leq \frac{h(t)}{f_1(t)} \leq \frac{1 + g(b(u))}{f_1(b(u)) - g(b(u))(t - b(u))}.$$

Now integrate between $b(u)$ and $a_1(u)x + b(u)$. Recall $a_1(u) = f_1(b(u))$ and

$$-\log u\bar{F}(a_1(u)x + b(u)) = \int_{b(u)}^{a_1(u)x+b(u)} \frac{h(t)}{f_1(t)}\, dt$$

and from (3.1) we get after a change of variable

$$\begin{aligned}(1 - g(b(u)))&(g(b(u)))^{-1}\log(1 + g(b(u))x)\\ &\leq -\log u\bar{F}(a_1(u)x + b(u))\\ &\leq -(1 + g(b(u)))(g(b(u)))^{-1}\log(1 - g(b(u))x).\end{aligned}$$

(The right inequality holds for $0 \leq x \leq 1/g(b(u))$ and the left for $x \geq 0$.) Exponentiating we find for $x \geq 0$

$$(\bar{F}(-g(b(u)), x))^{1+g(b(u))} \leq u\bar{F}(a_1(u)x + b(u)) \leq (\bar{F}(g(b(u)), x))^{1-g(b(u))}.$$

Since

$$\begin{aligned}\bar{F}(g, x)^{1-g} = (1 + gx)^{-(1-g)g^{-1}} &= \left(1 + \left(\frac{g}{1-g}\right)(x(1-g))\right)^{-(1-g)g^{-1}}\\ &= \bar{F}\left(\frac{g}{1-g}, x(1-g)\right)\end{aligned}$$

and

$$\bar{F}(-g, x)^{1+g} = \bar{F}(-(g/(1+g)), x(1+g))$$

we have

$$\begin{aligned}&\sup_{x\geq 0}|u\bar{F}(a_1(u)x + b(u)) - e^{-x}|\\ &\quad\leq \sup_{x\geq 0}\left|\bar{F}\left(\frac{g(b(u))}{1-g(b(u))}, x(1 - g(b(u)))\right) - e^{-x}\right|\\ &\quad\quad\vee \sup_{x\geq 0}\left|\bar{F}\left(\frac{-g(b(u))}{1+g(b(u))}, x(1 + g(b(u)))\right) - e^{-x}\right|.\end{aligned} \tag{3.2}$$

Now

$$\begin{aligned}\sup_{x\geq 0}&\left|\bar{F}\left(\frac{g(b(u))}{1-g(b(u))}, x(1 - g(b(u)))\right) - e^{-x}\right|\\ &\leq \sup_{x\geq 0}\left|\bar{F}\left(\frac{g(b(u))}{1-g(b(u))}, x(1 - g(b(u)))\right) - e^{-x(1-g(b(u)))}\right|\\ &\quad + \sup_{x\geq 0}|e^{-x(1-g(b(u)))} - e^{-x}|.\end{aligned}$$

According to Proposition 3.1, the first term is bounded by

$$\left(2+\frac{g(b(u))}{1-g(b(u))}\right)e^{-2}\frac{g(b(u))}{1-g(b(u))}$$

and using Proposition 3.2 (i), the second term is bounded by

$$\frac{g(b(u))}{1-g(b(u))}\,(1-g(b(u)))^{1/g(b(u))}$$

whence

$$\sup_{x\geq 0}\left|\bar{F}\left(\frac{g(b(u))}{1-g(b(u))},\,x(1-g(b(u)))\right)-e^{-x}\right|$$
$$\leq\left\{\left(2+\frac{g(b(u))}{1-g(b(u))}\right)e^{-2}+(1-g(b(u)))^{1/g(b(u))}\right\}\frac{g(b(u))}{1-g(b(u))}.$$

In a similar way using Proposition 3.1 and 3.2 (ii) we find

$$\sup_{x\geq 0}\left|\bar{F}\left(-\frac{g(b(u))}{1+g(b(u))},\,x(1+g(b(u)))\right)-e^{-x}\right|$$
$$\leq\left\{\left(2+\frac{g(b(u))}{1+g(b(u))}\right)e^{-2}+(1+g(b(u)))^{-1/g(b(u))}\right\}\frac{g(b(u))}{1+g(b(u))}.$$

Since $g/(1+g)\leq g/(1-g)$ and $(1+g)^{-1/g}\geq(1-g)^{1/g}$, we may bound the left side of (3.2) as follows:

$$\sup_{x\geq 0}|u\bar{F}(a_1(u)x+b(u))-e^{-x}|$$
$$\leq\left\{\left(2+\frac{g(b(u))}{1-g(b(u))}\right)e^{-2}+(1+g(b(u)))^{-1/g(b(u))}\right\}\frac{g(b(u))}{1-g(b(u))}\tag{3.3}$$
$$=\psi(g(b(u)))=O(g(b(u))).$$

REMARK. Reviewing the previous derivation we see that if $h\equiv 1$ then

$$\sup_{x\geq 0}|u\bar{F}(a_1(u)x+b(u))-e^{-x}|\leq(2+g(b(u)))e^{-2}g(b(u)).$$

The previous discussion contains the important ideas required for the following theorem on tail approximations.

THEOREM 3.1. *For $g>0$, let $\psi(g)$ be the function given by the right side of* (3.3); *i.e.*

$$\psi(g)=\left\{\left(2+\frac{g}{1-g}\right)e^{-2}+(1+g)^{-1/g}\right\}\frac{g}{1-g}.\tag{3.4}$$

Suppose F is a distribution in the domain of attraction of an extreme value distribution with $0<\alpha\leq\infty$.

(i) *If* $\alpha < \infty$ *and* F *satisfies* (2.1), (2.2), (2.3), *then for* $u > 1$

$$\sup_{s \ge b(u)} | u\bar{F}(s) - (s/b(u))^{-\alpha} | \le \psi(g_1(b(u))) \tag{3.5}$$

where g_1 *satisfies for* $t > 0$

$$g_1(t) \ge \alpha^{-1} | t/f_0(t) - \alpha |. \tag{3.6}$$

(ii) *If* $\alpha = \infty$ *and* F *satisfies* (2.6), *then for* $u > 1$

$$\sup_{s \ge b(u)} | u\bar{F}(s) - \exp\{-(s - b(u))/a_1(u)\} | \le \psi(g(b(u))) \tag{3.7}$$

where g *satisfies* (3.1). *Also*

$$\sup_{s \ge b(u)} | u\bar{F}(s) - (s/b(u))^{-b(u)/a_1(u)} | \le \psi(g_2(b(u))) \tag{3.8}$$

where g_2 *satisfies for* $t > 0$.

$$g_2(t) \ge | h(t) - 1 | \vee | f_1'(t) - t^{-1}f_1(t) |. \tag{3.9}$$

PROOF. The assertion (3.7) is just a restatement of (3.3). If we write (3.3) for the log transformed sample we get after changing variables ($y = a_1^* x + b^*$)

$$\sup_{y \ge b^*(u)} | u\bar{F}^*(y) - \exp\{-(y - b^*(u))/a_1^*(u)\} | \le \psi(g^*(b^*(u))). \tag{3.10}$$

If $\alpha < \infty$, then recalling $h^*(x) = \alpha^{-1}/(f_0(e^x)e^{-x})$, $f_1^*(x) \equiv \alpha^{-1}$ gives

$$g^*(x) \ge | h^*(x) - 1 | = \alpha^{-1} \left| \frac{e^x}{f_0(e^x)} - \alpha \right|.$$

Since $b^*(u) = \log b(u)$ we have

$$g^*(b^*(u)) \ge \alpha^{-1} \left| \frac{b(u)}{f_0(b(u))} - \alpha \right|.$$

Set $g^*(b^*(u)) = g_1(b(u))$ and recall $a_1^*(u) \equiv \alpha^{-1}$, so rewriting (3.10) by setting $s = e^y$ gives (3.5).

If $\alpha = \infty$, then $h^*(t) = h(e^t)$, $f_1^*(t) = f_1(e^t)/e^t$ so that $(f_1^*)'(t) = f_1'(e^t) - f_1(e^t)e^{-t}$. Therefore

$$g_2(b(u)) := g^*(b^*(u)) \ge | h(b(u)) - 1 | \vee | f_1'(b(u)) - f_1(b(u))/b(u) |$$

as asserted and setting $s = e^y$ in (3.10) and remembering $a_1^*(u) = a_1(u)/b(u)$ yields (3.8). □

We now discuss tail estimators. First note that when $\alpha = \infty$, b/a_1 in (3.8) is equal to $1/a_1^*$ (cf. 2.14) and that $a_1^*(t) \sim a^*(t) = t \int_{b^*(t)}^{\infty} \bar{F}^*(s)\, ds$. When $\alpha < \infty$, the exponent α in (3.5) is $1/a_1^*(t) \sim 1/a^*(t)$. Since $a^*(t)$ has a nice interpretation as the mean residual life of F^* evaluated at $b^*(t)$, we are led to consider what happens when a^*, b^* are replaced by empirical counterparts. As $b^*(n)$, the $1 - n^{-1}$ quantile of F^* is difficult to estimate from $X_1, \cdots, X_n$ (or $X_1^*, \cdots, X_n^*$), we consider instead $b^*(n/m)$, $a^*(n/m)$ where $m = m(n)$ satisfies $m/n \to 0$, $m \to \infty$. Thus we consider (3.5) and (3.8) with u replaced by n/m. In (3.5) we replace α, b by the estimators $1/\widehat{a^*}(n/m)$, $\hat{b}(n/m)$ and in (3.8), we replace b/a_1, b by $1/\widehat{a^*}(n/m)$, $\hat{b}(n/m)$.

THEOREM 3.2. *Suppose F is in the domain of attraction of an extreme value distribution as described in Section 2 with $0 < \alpha \le \infty$. Based on $X_1, \cdots, X_n$, suppose $\widehat{a^*} = \widehat{a^*}(n/m)$, $\hat{b} = \hat{b}(n/m)$ are any estimates of $a^*(n/m)$ and $b(n/m)$ satisfying ($n \to \infty$, $m \to \infty$, $n/m \to \infty$)*

$$\frac{\widehat{a^*}}{a^*} \to_P 1, \quad \frac{\widehat{b^*} - b^*}{a^*} := \frac{\log(\hat{b}/b)}{a^*} \to_P 0. \tag{3.11}$$

Then

$$\sup_{s \ge b(n/m)} |(n/m)\bar{F}(s) - (s/\hat{b})^{-1/\widehat{a^*}}| \to_P 0 \tag{3.12}$$

so that $(s/\hat{b})^{-1/\widehat{a^*}}$ is a consistent tail estimate for $\bar{F}(s)/\bar{F}(b(n/m))$, $s \ge b(n/m)$.

PROOF. The $\alpha = \infty$ case is slightly harder than $\alpha < \infty$ so we concentrate on $\alpha = \infty$. Write

$$\begin{aligned}\sup_{s \ge b} &|(n/m)\bar{F}(s) - (s/\hat{b})^{-1/\widehat{a^*}}| \\ &\le \sup_{s \ge b}|(n/m)\bar{F}(s) - (s/b)^{-1/a_1^*}| + \sup_{s \ge b}|(s/b)^{-1/a_1^*} - (s/b)^{-1/\widehat{a^*}}| \\ &\quad + \sup_{s \ge b}|(s/b)^{-1/\widehat{a^*}} - (s/\hat{b})^{-1/\widehat{a^*}}| = A + B + C.\end{aligned}$$

For A, apply (3.8) to get

$$A \le O(g_2(b(n/m))) \to 0 \quad \text{as} \quad n \to \infty.$$

Next we have

$$B = \sup_{y \ge 1}|y^{-1/a_1^*} - y^{-1/\widehat{a^*}}|$$

which from Proposition 3.2 (iii) is bounded by

$$e^{-1}\left|\frac{1/a_1^* - 1/\widehat{a^*}}{1/a_1^* \wedge 1/\widehat{a^*}}\right| = e^{-1}\left|\frac{\widehat{a^*}/a_1^* - 1}{\widehat{a^*}/a_1^* \wedge 1}\right| \to_P 0$$

because of (3.11) and the fact that $a^* \sim a_1^*$. Finally we have that

$$C = b^{-1/\widehat{a^*}}|b^{1/\widehat{a^*}} - \hat{b}^{1/\widehat{a^*}}| = |(\hat{b}/b)^{1/\widehat{a^*}} - 1| = |\exp\{(\widehat{b^*} - b^*)/\widehat{a^*}\} - 1|.$$

Since $\widehat{a^*}/a^* \to_P 1$ we have $(\widehat{b^*} - b^*)/\widehat{a^*} \to_P 0$ whence

$$|\exp\{(\widehat{b^*} - b^*)/\widehat{a^*}\} - 1| \to_P 0$$

as required. □

REMARK. Strong consistency in (3.11) leads to strong consistency in (3.12).

We now specify the estimators $\widehat{a^*}$, $\hat{b}$. Recalling that $X_{(1)} > X_{(2)} >, \cdots, > X_{(n)}$ denote the decreasing order statistics of the sample $X_1, \cdots, X_n$ (similar notation for the log transformed sample) we set

$$\hat{b}(n/m) = X_{(m+1)}$$

the empirical $1 - m/n$ quantile and

$$\widehat{a^*}(n/m) = m^{-1} \sum_{i=1}^{m} (X^*_{(i)} - X^*_{(m+1)}) = m^{-1} \sum_{i=1}^{m} (\log X_{(i)} - \log X_{(m+1)})$$

the empirical mean residual life of the transformed data evaluated at $X^*_{(m+1)}$. The asymptotic properties of these estimators are studied in the next sections and in particular we check that (3.11) is valid.

4. Asymptotic propertis of â, b̂, and $\hat{\alpha}$. In this section we consider the asymptotic properties of the estimates of $a(n/m) = n/m \int_{b(n/m)}^{\infty} \bar{F}(s)\, ds$ and $b(n/m) = F^{\leftarrow}(1 - m/n)$ where $\bar{F}$ satisfies (2.1) and either (2.3) with $\alpha > 2$, or (2.6). The estimates are given by

$$\hat{a}(n/m) = m^{-1} \sum_{j=1}^{m} (X_{j,n} - X_{m+1,n}) \tag{4.1}$$

and

$$\hat{b}(n/m) = X_{m+1,n} \tag{4.2}$$

where $X_{1,n} \geq X_{2,n} \geq \cdots \geq X_{n,n}$ are the decreasing order statistics from a random sample $X_1, \cdots, X_n$ with df. F.Via the probability integral transform, the random variables defined by $E_j = -\log(1 - F(X_j))$ are independent unit exponentials. Since $F(X_j) = 1 - e^{-E_j}$, we have $b(e^{E_j}) = X_j$ a.s. So if $E_{1,n} \geq E_{2,n}, \cdots, \geq E_{n,n}$, denote the decreasing order statistics based on $E_1, \cdots, E_n$, then with probability one,

$$\hat{a}(n/m) = m^{-1} \sum_{j=1}^{m} (b(e^{E_{j,n}}) - b(e^{E_{m+1,n}})) \tag{4.3}$$

and

$$\hat{b}(n/m) = b(e^{E_{m+1,n}}). \tag{4.4}$$

The proofs of the following results rely on this representation and the well known properties of the order statistics from an exponential sample.

THEOREM 4.1. *Suppose F belongs to the α domain of attraction with $2 < \alpha \leq \infty$. Then as $n \to \infty$*

$$\begin{aligned} &\left(\sqrt{m}\left(\frac{\hat{a}(n/m)}{a(n/m)} - 1\right), \sqrt{m}\left(\frac{\hat{b}(n/m) - b(n/m)}{a(n/m)}\right)\right) \\ &\qquad \Rightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha^{-2} + \alpha/(\alpha - 2) & (\alpha - 1)/\alpha^2 \\ (\alpha - 1)/\alpha^2 & ((\alpha - 1)/\alpha)^2 \end{bmatrix}\right) \end{aligned} \tag{4.5}$$

where $m = m(n)$ is any sequence of integers satisfying $m \to \infty$ and $m/n \to 0$. Moreover,

$$\frac{\hat{a}(n/m)}{a(n/m)} \to_P 1, \quad \frac{\hat{b}(n/m) - b(n/m)}{a(n/m)} \to_P 0, \quad \text{and} \quad \frac{\hat{a}(n/m)}{\hat{a}(n/m) + \hat{b}(n/m)} \to_P \alpha^{-1}.$$

REMARK. The formulas in Theorem 4.1 in the $\alpha = \infty$ case are to be interpreted as the limit $\alpha \to \infty$. Notice that in this case the two components in (4.5) are asymptotically independent.

The proof of this theorem will be broken up into a series of lemmas.

LEMMA 4.1. *For any $\varepsilon > 0$ with $\varepsilon < (\alpha - 1)/\alpha$, there exists a positive real number T such that*

$$\left| \frac{a_0(ts)}{a(t)} - \frac{\alpha - 1}{\alpha} s^{\alpha^{-1}} \right| \le \left(\frac{\alpha - 1}{\alpha} + \varepsilon\right) s^{\alpha^{-1}+\varepsilon} - \frac{\alpha - 1}{\alpha} s^{\alpha^{-1}} \tag{4.6}$$

for all $t \ge T$ and $s \ge 1$ where $a_0(t)$ is given by (2.4). Furthermore $a_0(ts)/a(t) \to ((\alpha - 1)/\alpha)s^{\alpha^{-1}}$ uniformly on a compact neighborhood of $s = 1$ as $t \to \infty$.

PROOF. Define

$$u(t) = \frac{ta'(t)}{a(t)} = \frac{-t^2\bar{F}(b(t))b'(t) + a(t)}{a(t)} = 1 - \frac{a_0(t)}{a(t)} \tag{4.7}$$

where we have used (2.10). Since $a_0(t)/a(t) \to (\alpha - 1)/\alpha$, we have $u(t) \to 1/\alpha$ and hence for any $0 < \varepsilon < (\alpha - 1)/\alpha$, there exists a positive number T such that

$$\alpha^{-1} - \varepsilon \le u(t) \le \alpha^{-1} + \varepsilon \quad \text{for all} \quad t \ge T. \tag{4.8}$$

Solving the differential equation (4.7) for $a(t)$, we obtain

$$a(t) = a(1)\exp\left(\int_1^t \frac{u(x)}{x\,dx}\right) \tag{4.9}$$

whence

$$a(ts)/a(t) = \exp\left(\int_t^{ts} \frac{u(x)}{x}\,dx\right) = \exp\left(\int_1^s \frac{u(tx)}{x}\,dx\right) \quad \text{for all} \quad s \ge 1.$$

Applying the bounds in (4.8) to the exponent we get (cf. Feller, 1971, page 277)

$$s^{\alpha^{-1}-\varepsilon} \le a(ts)/a(t) \le s^{\alpha^{-1}+\varepsilon} \quad \text{for all} \quad t \ge T \quad \text{and} \quad s \ge 1. \tag{4.10}$$

Moreover the inequality (4.8) also gives the bounds

$$\frac{\alpha - 1}{\alpha} - \varepsilon \le \frac{a_0(t)}{a(t)} \le \frac{\alpha - 1}{\alpha} + \varepsilon \quad \text{for all} \quad t \ge T. \tag{4.11}$$

Writing $a_0(ts)/a(t) = (a(ts)/a(t))(a_0(ts)/a(ts))$ the inequality

$$\left(\frac{\alpha - 1}{\alpha} - \varepsilon\right) s^{\alpha^{-1}-\varepsilon} \le \frac{a_0(ts)}{a(t)} \le s^{\alpha^{-1}+\varepsilon}\left(\frac{\alpha - 1}{\alpha} + \varepsilon\right) \tag{4.12}$$

for all $t \ge T$ and $s \ge 1$ follows from (4.10) and (4.11). Now subtracting $((\alpha - 1)/\alpha)s^{\alpha^{-1}}$ from each side in (4.12), the bound (4.6) follows easily.

For $s < 1$, the inequality (4.10) is reversed for $t > T$ so that (4.12) becomes

$$\left(\frac{\alpha-1}{\alpha}-\varepsilon\right)s^{\alpha^{-1}+\varepsilon} \le \frac{a_0(ts)}{a(t)} \le s^{\alpha^{-1}-\varepsilon}\left(\frac{\alpha-1}{\alpha}+\varepsilon\right) \quad \text{for} \quad 0 < s < 1.$$

It is then clear that $a_0(ts)/a(t) \to ((\alpha - 1)/\alpha)s^{\alpha^{-1}}$ uniformly on a compact neighborhood of $s = 1$.

LEMMA 4.2. *For* $2 < \alpha \le \infty$,

$$\left(\sqrt{m}\left(\frac{\hat{b}(n/m) - b(n/m)}{a(n/m)}\right), \sqrt{m}\left(\frac{a(e^{E_{m+1,n}})}{a(n/m)} - 1\right)\right) \Rightarrow \left(\frac{\alpha-1}{\alpha}, \frac{1}{\alpha}\right) \cdot Z_1$$

where Z_1 *is a* $N(0, 1)$ *random variable.*

PROOF. First, from Renyi's representation for order statistics from an exponential sample (cf. Feller, 1971, page 19) we have

$$E_{m+1,n} =_d \textstyle\sum_{j=1}^{n-m} (n + 1 - j)^{-1}E_j.$$

It follows by checking characteristic functions directly (or see Smirnov, 1952, Theorem 4) that

$$\sqrt{m}(E_{m+1,n} - \log(n/m)) \Rightarrow Z_1$$

where Z_1 is a $N(0, 1)$ random variable. By appealing to Skorokhod's theorem (Billingsley, 1979), we shall assume for the remainder of the proof that

$$N_n: = \sqrt{m}(E_{m+1,n} - \log(n/m)) \to Z_1 \quad \text{a.s.} \tag{4.13}$$

Now since $b'(t) = a_0(t)/t$ (see (2.10)), we have from (4.4) and (4.13)

$$\sqrt{m}\left(\frac{\hat{b}(n/m) - b(n/m)}{a(n/m)}\right)$$
$$= \sqrt{m}\int_{n/m}^{\exp(E_{m+1,n})} \frac{a_0(s)}{a(n/m)s}\,ds = \sqrt{m}\int_1^{\exp(N_n/\sqrt{m})} \frac{a_0((n/m)s)}{a(n/m)s}\,ds.$$

By Lemma 4.1,

$$(a_0((n/m)s)/a(n/m)) \to ((\alpha - 1)/\alpha)s^{1/\alpha}$$

uniformly on a compact neighborhood of 1 and since $N_n/\sqrt{m} \to 0$ a.s. it follows that

$$\sqrt{m}\int_1^{\exp(N_n/\sqrt{m})} \frac{a_0((n/m)s)}{a(n/m)s}\,ds$$
$$= \sqrt{m}(\alpha - 1)(e^{\alpha^{-1}N_n/\sqrt{m}} - 1) + o(1) \to \frac{\alpha-1}{\alpha} \cdot Z_1 \quad \text{a.s.}$$

as desired.

From the representation (4.9), we have

$$\frac{a(e^{E_{m+1,n}})}{a(n/m)} = \exp\left(\int_{n/m}^{\exp(E_{m+1,n})} \frac{u(x)}{x}\,dx\right) = \exp\left(\int_{1}^{\exp(N_n/\sqrt{m})} \frac{u((n/m)x)}{x}\,dx\right)$$

where $u(t) \to 1/\alpha$ as $t \to \infty$. Consequently $u((n/m)x) \to 1/\alpha$ uniformly on a compact neighborhood of 1 so that (4.13) implies

$$\sqrt{m}\left(\frac{a(e^{E_{m+1,n}})}{a(n/m)} - 1\right) = \sqrt{m}(e^{\alpha^{-1}N_n/\sqrt{m}} - 1) + o(1) \to \alpha^{-1}Z_1 \quad \text{a.s.}$$

which completes the proof. □

LEMMA 4.3. *For* $2 < \alpha \le \infty$,

$$\sqrt{m}((\hat{a}(n/m)/a(e^{E_{m+1,n}})) - 1) \Rightarrow (\alpha/(\alpha-2))^{1/2}Z_2$$

where Z_2 *is a* $N(0, 1)$ *random variable independent of* Z_1 *in Lemma* 4.2.

PROOF. Observe that for a random sample of exponential random variables

$$(E_{1,n} - E_{m+1,n}, E_{2,n} - E_{m+1,n}, \cdots, E_{m,n} - E_{m+1,n}, E_{m+1,n}) =_d (E_{1,m}, E_{2,m}, \cdots, E_{m,m}, E_{m+1,n})$$

where $E_{m+1,n}$ is independent of $(E_{1,m}, \cdots, E_{m,m})$. Thus from (4.3)

$$\begin{aligned}\hat{a}(n/m) &= m^{-1}\textstyle\sum_{j=1}^m (b(e^{E_{j,n}}) - b(e^{E_{m+1,n}}))\\ &= m^{-1}\textstyle\sum_{j=1}^m (b(e^{E_{m+1,n}}e^{E_{j,n}-E_{m+1,n}}) - b(e^{E_{m+1,n}}))\\ &=_d m^{-1}\textstyle\sum_{j=1}^m (b(e^{E_{m+1,n}+E_{j,m}}) - b(e^{E_{m+1,n}}))\end{aligned}$$

and using symmetry we have

$$\hat{a}(n/m) =_d m^{-1}\textstyle\sum_{j=1}^m (b(e^{E_{m+1,n}+E_j}) - b(e^{E_{m+1,n}})) \tag{4.14}$$

where for all n, $E_{m+1,n}$ is independent of the sequence of independent unit exponential random variables $(E_1, E_2, \cdots)$. It follows from (4.14) that

$$\begin{aligned}E(\hat{a}(n/m) \mid E_{m+1,n}) &= \int_0^\infty (b(e^{E_{m+1,n}+x}) - b(e^{E_{m+1,n}}))e^{-x}\,dx\\ &= e^{E_{m+1,n}}\int_{E_{m+1,n}}^\infty b(e^x)e^{-x}\,dx - b(e^{E_{m+1,n}}).\end{aligned} \tag{4.15}$$

However, since

$$\begin{aligned}a(t) &= t\int_{b(t)}^\infty \bar{F}(y)\,dy = t\int_{b(t)}^\infty xF(dx) - b(t)\\ &= t\int_{\log t}^\infty b(e^x)e^{-x}\,dx - b(t),\end{aligned}$$

we conclude that (4.15) is equal to $a(e^{E_{m+1,n}})$ and therefore

$$E(\hat{a}(n/m)/a(e^{E_{m+1,n}}) \mid E_{m+1,n}) = 1 \quad \text{a.s.} \tag{4.16}$$

We rewrite (4.14) as follows:

$$\frac{\hat{a}(n/m)}{a(e^{E_{m+1,n}})} =_d m^{-1} \sum_{j=1}^m \left(\frac{b(e^{E_{m+1,n}+E_j}) - b(e^{E_{m+1,n}})}{a(e^{E_{m+1,n}})} - (\alpha - 1)(e^{\alpha^{-1}E_j} - 1) \right) + (\alpha - 1)m^{-1} \sum_{j=1}^m (e^{\alpha^{-1}E_j} - 1)$$

(which is suggested by the limit relation $\lim_{t\to\infty}(b(tx) - b(t))/a(t) = \lim_{t\to\infty} \int_t^{tx} (b'(s)/a(t))\, ds = \lim_{t\to\infty} \int_1^x (a_0(ts)/a(t))(ds/s) = (\alpha - 1)(x^{\alpha^{-1}} - 1)$ for all $1 < \alpha \le \infty$; see Lemma 4.1). Therefore,

$$\sqrt{m}\left(\frac{\hat{a}(n/m)}{a(e^{E_{m+1,n}})} - 1\right) =_d \frac{1}{\sqrt{m}} \sum_{j=1}^m Y_n^{(j)} + \frac{\alpha - 1}{\sqrt{m}} \sum_{j=1}^m \left(e^{\alpha^{-1}E_j} - \frac{\alpha}{\alpha - 1}\right) \tag{4.17}$$

where using the fact that $b'(s) = a_0(s)/s$ we may write

$$Y_n^{(j)} = \int_1^{\exp(E_j)} \left(\frac{a_0(e^{E_{m+1,n}}s)}{a(e^{E_{m+1,n}})} - \left(\frac{\alpha - 1}{\alpha}\right) s^{\alpha^{-1}} \right) \frac{ds}{s}.$$

Since $Ee^{\alpha^{-1}E_j} = \alpha/(\alpha - 1)$ and $\mathrm{Var}(e^{\alpha^{-1}E_j}) = \alpha/((\alpha - 1)^2(\alpha - 2))$, the second expression on the right of (4.17) converges in distribution to $(\alpha/(\alpha - 2))^{1/2}Z_2$ where $Z_2 \sim N(0, 1)$ and is independent of the random variable Z_1 of Lemma 4.2 by the independence of the E_j's and $E_{m+1,n}$.

To complete the proof we will show $\sum_{j=1}^m Y_n^{(j)}/\sqrt{m} \to_p 0$ and for this it suffices to show for every $\delta > 0$ that $P[\,| m^{-1/2} \sum_{j=1}^m Y_n^{(j)} | > \delta \mid E_{m+1,n}] \to_p 0$. Observe that $E(Y_n^{(j)} \mid E_{m+1,n}) = 0$ for all $j, n, j \le m$ by (4.16) and the independence of $E_{m+1,n}$ and the E_j's. Therefore by Chebychev

$$\begin{aligned} &P[\,| m^{-1/2} \textstyle\sum_{j=1}^m Y_n^{(j)} | > \delta \mid E_{m+1,n}] \\ &\qquad \le 1/\delta^2 E(| m^{-1/2} \textstyle\sum_{j=1}^m Y_n^{(j)} |^2 \mid E_{m+1,n}) 1_{[E_{m+1,n}>T]} \\ &\qquad\quad + P(| m^{-1/2} \textstyle\sum_{j=1}^m Y_n^{(j)} | > \delta \mid E_{m+1,n}) 1_{[E_{m+1,n}\le T]}. \end{aligned} \tag{4.18}$$

Since for each n, $Y_n^{(1)}, Y_n^{(2)}, \cdots$ are conditionally independent and identically distributed given $E_{m+1,n}$ with conditional mean 0, the first term in (4.18) is equal to $\delta^{-2}E((Y_n^{(1)})^2 \mid E_{m+1,n})1_{[E_{m+1,n}>T]}$ a.s. Applying the inequality (4.6) to the integrand of $Y_n^{(1)}$, we obtain

$$\begin{aligned} &E((Y_n^{(1)})^2 \mid E_{m+1,n}) 1_{[E_{m+1,n}>T]} \\ &\le E\left[\left(\frac{\alpha - 1}{\alpha} + \varepsilon\right)\left(\frac{1}{\alpha} + \varepsilon\right)^{-1} (e^{(\alpha^{-1}+\varepsilon)E_1} - 1) - (\alpha - 1)(e^{\alpha^{-1}E_1} - 1) \right]^2 \cdot 1_{[E_{m+1,n}>T]}. \end{aligned}$$

The expectation portion of this bound does not depend on n and approaches 0 as $\varepsilon \to 0$ by the dominated convergence theorem. Hence this term goes to zero in probability as $n \to \infty$. As for the second term in (4.18), it also goes to zero in probability since $E_{m+1,n} \to \infty$ in probability by (4.13), which finishes the proof. □

PROOF OF THEOREM 4.1. The independence of Z_1 and Z_2 in Lemmas 4.2 and 4.3 implies that

$$\left(\sqrt{m}\,\frac{\hat{b}(n/m) - b(n/m)}{a(n/m)},\ \sqrt{m}\left(\frac{a(e^{E_{m+1,n}})}{a(n/m)} - 1\right),\ \sqrt{m}\left(\frac{\hat{a}(n/m)}{a(e^{E_{m+1,n}})} - 1\right)\right)$$
$$\Rightarrow (\alpha^{-1}(\alpha - 1)Z_1,\ \alpha^{-1}Z_1,\ (\alpha/(\alpha - 2))^{1/2}Z_2).$$

Since

$$\sqrt{m}\left(\frac{\hat{a}(n/m)}{a(n/m)} - 1\right) = \sqrt{m}\left(\frac{\hat{a}(n/m)}{a(e^{E_{m+1,n}})} - 1\right) + \frac{\hat{a}(n/m)}{a(e^{E_{m+1,n}})}\sqrt{m}\left(\frac{a(e^{E_{m+1,n}})}{a(n/m)} - 1\right)$$

and $\hat{a}(n/m)/a(e^{E_{m+1,n}}) \to 1$ in probability by Lemma 4.3, it follows that

$$\left(\sqrt{m}\left(\frac{\hat{a}(n/m)}{a(n/m)} - 1\right),\ \sqrt{m}\left(\frac{\hat{b}(n/m) - b(n/m)}{a(n/m)}\right)\right)$$
$$\Rightarrow (\alpha^{-1}Z_1 + (\alpha/(\alpha - 2))^{1/2}Z_2,\ \alpha^{-1}(\alpha - 1)Z_1) \tag{4.19}$$

which establishes (4.5).

The convergence in probability of $\hat{a}(n/m)/a(n/m)$ and $(\hat{b}(n/m) - b(n/m))/a(n/m)$ to 1 and 0 respectively is immediate from (4.5). This implies that $(\hat{b}(n/m) - b(n/m))/\hat{a}(n/m) \to_P 0$ and since $b(n/m)/a(n/m) \to \alpha - 1$, we must have $b(n/m)/\hat{a}(n/m) = (a(n/m)/\hat{a}(n/m))(b(n/m)/a(n/m)) \to_P \alpha - 1$. Hence, $\hat{b}(n/m)/\hat{a}(n/m) \to_P \alpha - 1$ which yields the result

$$\hat{a}(n/m)/(\hat{a}(n/m) + \hat{b}(n/m)) \to_P 1/\alpha. \quad \square$$

We now apply Theorem 4.1 to the log transformed variables discussed in Section 2. Recall that $F^*(x) = F(e^x)$ denotes the distribution of $\log X_1$ and F^* is in the $\alpha = \infty$ domain of attraction. The corresponding estimates of $a^*(n/m) = n/m \int_{\log b(n/m)}^{\infty} \bar{F}^*(s)\, ds$ and $b^*(n/m) = \log b(n/m)$ are

$$\widehat{a^*}(n/m) = m^{-1} \textstyle\sum_{j=1}^{m} (\log X_{j,n} - \log X_{m+1,n})$$

and

$$\widehat{b^*}(n/m) = \log X_{m+1,n} = \log \hat{b}(n/m).$$

The asymptotic behavior of these two estimates is immediate from Theorem 4.1.

COROLLARY 1. *Suppose F belongs to the α-domain of attraction with $0 < \alpha \le \infty$. Then*

$$\left(\sqrt{m}\left(\frac{\widehat{a^*}(n/m)}{a^*(n/m)} - 1\right),\ \sqrt{m}\left(\frac{\widehat{b^*}(n/m) - b^*(n/m)}{a^*(n/m)}\right)\right) \Rightarrow N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right).$$

In particular

$$\widehat{a^*}(n/m)/a^*(n/m) \to_P 1,\ (\widehat{b^*}(n/m) - b^*(n/m))/a^*(n/m) \to_P 0,$$

and

$$\widehat{a^*}(n/m) \to_P 1/\alpha.$$

REMARK. Since $a^*(m/n) \to \alpha^{-1}$, it follows in the $\alpha < \infty$ case by the convergence to types theorem, that if for some choice of $m \to \infty$, $\sqrt{m}(a^*(n/m) - \alpha^{-1}) \to c$, then

$$\sqrt{m}(\widehat{a^*}(n/m) - \alpha^{-1}) \Rightarrow N(c, \alpha^{-2}).$$

However the condition for this to hold depends heavily on the unknowns F and α and hence this result cannot be used for obtaining a confidence interval for α (cf. Hall, 1982).

Another weakly consistent estimate of α^{-1} for $\alpha > 2$ is

$$\hat{a}(n/m)/(\hat{a}(n/m) + \hat{b}(n/m)).$$

However as shown in the following result, this estimate has a larger asymptotic variance than that of $\widehat{a^*}(n/m)$.

COROLLARY 2. *For* $2 < \alpha \le \infty$

$$\sqrt{m}\left(\frac{\hat{a}(n/m)}{\hat{a}(n/m) + \hat{b}(n/m)} \cdot \frac{a(n/m) + b(n/m)}{a(n/m)} - 1\right) \Rightarrow N\left(0, \frac{(\alpha - 1)^2}{\alpha(\alpha - 2)}\right).$$

PROOF. We have

$$\begin{aligned} &\sqrt{m}\left(\frac{\hat{a}(n/m)}{\hat{a}(n/m) + \hat{b}(n/m)} \cdot \frac{a(n/m) + b(n/m)}{a(n/m)} - 1\right) \\ &= \frac{b(n/m)}{\hat{a}(n/m) + \hat{b}(n/m)} \sqrt{m}\left(\frac{\hat{a}(n/m)}{a(n/m)} - 1\right) \\ &\quad - \frac{a(n/m)}{\hat{a}(n/m) + \hat{b}(n/m)} \sqrt{m}\, \frac{\hat{b}(n/m) - b(n/m)}{a(n/m)}. \end{aligned} \tag{4.20}$$

Since $\hat{b}(n/m)/b(n/m) \to_P 1$ and $\hat{a}(n/m)/a(n/m) \to_P 1$ by Theorem 4.1, it follows that $b(n/m)/(\hat{a}(n/m) + \hat{b}(n/m)) \to_P (\alpha - 1)/\alpha$ and $a(n/m)/(\hat{a}(n/m) + \hat{b}(n/m)) \to_P 1/\alpha$ and hence the right-hand side of (4.20) converges in distribution to

$$\begin{aligned} &\frac{\alpha - 1}{\alpha}\left(\frac{1}{\alpha} Z_1 + \left(\frac{\alpha}{\alpha - 2}\right)^{1/2} Z_2\right) - \frac{1}{\alpha}\left(\frac{\alpha - 1}{\alpha}\right) Z_1 \\ &\qquad = \frac{\alpha - 1}{\alpha}\left(\frac{\alpha}{\alpha - 2}\right)^{1/2} Z_2 \sim N\left(0, \frac{(\alpha - 1)^2}{\alpha(\alpha - 2)}\right) \end{aligned}$$

by (4.19)

REMARK. The estimator $\hat{a}/(\hat{a} + \hat{b})$ is approximately normal with mean

$a/(a+b) \sim \alpha^{-1}$ and variance

$$m^{-1}\frac{(\alpha-1)^2}{\alpha(\alpha-2)}\left(\frac{a}{a+b}\right)^2 \sim m^{-1}\frac{(\alpha-1)}{\alpha^3(\alpha-2)}$$

(assuming $2 < \alpha < \infty$). The estimator $\widehat{a^*}$ is approximately normal with mean $a^* \sim \alpha^{-1}$ and variance

$$m^{-1}(a^*)^2 \sim m^{-1}\alpha^{-2} < m^{-1}\frac{(\alpha-1)^2}{\alpha^3(\alpha-2)}.$$

An additional advantage of $\widehat{a^*}$ is that the asymptotic distribution theory is not dependent on $\alpha > 2$ as is the case with $\hat{a}/(\hat{a}+\hat{b})$.

5. Strong consistency. The weak consistency results of Theorem 4.1 and Corollary 1 may be strengthened to strong consistency for special choices of the $m(n)$ sequence. These results in the $\alpha = \infty$ case partially extend those of Mason (1982). Also see Teugels (1981) and Haeusler and Schneemeirer (1983). We begin with two preparatory lemmas.

LEMMA 5.1. *Let $E_1, E_2, \cdots$ be an i.i.d. sequence of unit exponentials with $E_{1,n} \ge E_{2,n} \ge \cdots \ge E_{n,n}$ denoting the decreasing order statistics from the first n. If $m = m(n) = [n^\delta]$, $0 < \delta < 1$, ($[x]$ denotes the greatest integer $\le x$), then*

$$E_{m+1,n} - \log(n/m) \to 0 \quad a.s. \text{ as } \quad n \to \infty. \tag{5.1}$$

PROOF. Given $\varepsilon > 0$,

$$P[E_{m+1,n} > \varepsilon + \log(n/m)] = P[S_n > m]$$

where $S_n = \#\{1 \le j \le n: E_j > \varepsilon + \log(n/m)\}$ has a binomial (n, p_n) distribution with $p_n = P[E_1 > \varepsilon + \log(n/m)] = m/n\, e^{-\varepsilon}$. Since $m - np_n = m(1-e^{-\varepsilon}), > 0$, we have by Bernstein's inequality (Serfling, 1980, page 95)

$$\begin{aligned} P[S_n > m] &= P[S_n - np_n > m - np_n] \le P[|S_n - np_n| > m - np_n] \\ &\le 2\exp\{-n((m-np_n)/n)^2/(2(p_n + (m-np_n)/n))\} \\ &= 2\exp\{-m(1-e^{-\varepsilon})^2/2\} \\ &= 2\exp\{-[n^\delta](1-e^{-\varepsilon})^2/2\}. \end{aligned}$$

It follows that

$$\textstyle\sum_{n=1}^\infty P[E_{m+1,n} - \log(n/m) > \varepsilon] < \infty.$$

The proof will be complete by an application of Borel-Cantelli, once we show

$$\textstyle\sum_{n=1}^\infty P[E_{m+1,n} - \log(n/m) < -\varepsilon] < \infty. \tag{5.2}$$

However

$$P[E_{m+1,n} - \log(n/m) < -\varepsilon] = P[\tilde{S}_n < m]$$

where $\tilde{S}_n = \#\{j: E_j > -\varepsilon + \log(n/m)\}$ has a binomial (n, p_n) distribution with

$p_n = (m/n)e^{\epsilon}$. Following the same argument as above, we obtain the bound $P[\tilde{S}_n < m] \le 2\exp\{-[n^{\delta}](e^{\epsilon} - 1)/(4e^{\epsilon} - 2)\}$ which establishes (5.2) as desired. □

LEMMA 5.2. *Under the hypotheses of Lemma* (5.1), *we have for any* $1 < \alpha \le \infty$,

$$(5.3) \qquad (\alpha - 1)^{m^{-1}} \textstyle\sum_{j=1}^{m} (e^{\alpha^{-1}(E_{j,n} - E_{m+1,n})} - 1) \to 1 \quad \textit{a.s. as} \quad n \to \infty.$$

PROOF. For the $\alpha = \infty$ case, the left-hand side of (5.3) is

$$m^{-1} \textstyle\sum_{j=1}^{m} (E_{j,n} - E_{m+1,n})$$

and the result follows immediately from Theorem 1 in Mason (1982). So assume $1 < \alpha < \infty$ and let $G(x) = P[e^{\alpha^{-1}E_1} \le x] = 1 - x^{-\alpha}$. If $G_n(x)$ denotes the empirical distribution function of $e^{\alpha^{-1}E_1}, \dots, e^{\alpha^{-1}E_n}$, then

$$\begin{aligned}(\alpha - 1)m^{-1} &\textstyle\sum_{j=1}^{m} (e^{\alpha^{-1}(E_{j,n} - E_{m+1,n})} - 1) \\ &= (\alpha - 1)n/m \int_{\exp(\alpha^{-1}E_{m+1,n})}^{\infty} \bar{G}_n(x)\, dx \cdot e^{-\alpha^{-1}E_{m+1,n}}.\end{aligned}$$

Since

$$(\alpha - 1)n/m \int_{\exp(\alpha^{-1}E_{m+1,n})}^{\infty} \bar{G}(x)\, dx \cdot e^{-\alpha^{-1}E_{m+1,n}} = e^{-(E_{m+1,n} - \log(n/m))} \to 1 \quad \text{a.s.}$$

by Lemma 5.1, it is enough to show

$$(\alpha - 1)e^{-\alpha^{-1}E_{m+1,n}}(n/m) \int_{\exp(\alpha^{-1}E_{m+1,n})}^{\infty} |\bar{G}_n(x) - \bar{G}(x)|\, dx \to 0 \quad \text{a.s.}$$

But the left-hand side is equal to

$$\begin{aligned}(\alpha - 1)e^{-\alpha^{-1}E_{m+1,n}}(n/m) &\int_{\exp(\alpha^{-1}E_{m+1,n})}^{\infty} \frac{|\bar{G}_n(x) - \bar{G}(x)|}{\bar{G}(x)^{(\alpha+1)/(2\alpha)}} \bar{G}(x)^{(\alpha+1)/(2\alpha)}\, dx \\ &\le e^{-\alpha^{-1}E_{m+1,n}}(n/m)\sup_{x>0} \frac{|\bar{G}_n(x) - \bar{G}(x)|}{\bar{G}(x)^{(\alpha+1)/(2\alpha)}} 2e^{-(\alpha-1)/(2\alpha)E_{m+1,n}} \\ &\le 2e^{-(\alpha+1)/(2\alpha)(E_{m+1,n} - \log(n/m))}(n/m)^{(\alpha-1)/(2\alpha)}\sup_{0<t<1} \frac{|G_n(G^{-1}(t)) - t|}{(t(1-t))^{(\alpha+1)/(2\alpha)}}\end{aligned}$$

which, by Lemma 5.1, is asymptotically equivalent to

$$(5.4) \qquad (n/m)^{(\alpha-1)/(2\alpha)} \sup_{0<t<1} \frac{|G_n(G^{-1}(t)) - t|}{(t(1-t))^{(\alpha+1)/(2\alpha)}}.$$

However since $n/m \sim n^{1-\delta}$ and $G_n(G^{-1}(t))$ is the empirical df for a random sample from a uniform (0, 1) distribution, we may apply Corollary 1 of Mason (1981) with $v = (\alpha - 1)/(2\alpha)$ to show that (5.4) goes to zero almost surely. □

THEOREM 5.1. *Suppose* F *belongs to the* α*-domain of attraction with* $1 < \alpha \le$

∞. *Then if* $m = [n^\delta]$ *with* $0 < \delta < 1$,

(i) $\dfrac{\hat{b}(n/m) - b(n/m)}{a(n/m)} \to 0$ *a.s.*

(ii) $\dfrac{\hat{a}(n/m)}{a(n/m)} \to 1$ *a.s.*

(iii) $\dfrac{\hat{a}(n/m)}{\hat{a}(n/m) + \hat{b}(n/m)} \to 1/\alpha$ *a.s.*

(iv) $\widehat{a^*}(n/m) \to 1/\alpha$ *a.s. for all* $\alpha > 0$.

REMARK. Under the conditions of Theorem 5.1, the tail estimate given in Theorem 3.2 is strongly consistent.

PROOF. (i) From (2.10), as in the proof of Lemma 4.2, we have

$$\frac{\hat{b}(n/m) - b(n/m)}{a(n/m)} = \int_1^{\exp(E_{m+1,n} - \log(n/m))} \frac{a_0((n/m)s)}{a(n/m)s}\, ds$$

and since $a_0((n/m)s)/a(n/m) \to ((\alpha - 1)/\alpha)s^{\alpha^{-1}}$ uniformly on a compact neighborhood of 1, the result follows from Lemma 5.1.

(ii) From (4.3) and (2.10) we have

$$\frac{\hat{a}(n/m)}{a(e^{E_{m+1,n}})} = m^{-1} \sum_{j=1}^m \int_1^{\exp(E_{j,n} - E_{m+1,n})} \frac{a_0(e^{E_{m+1,n}}s)}{a(e^{E_{m+1,n}})s}\, ds$$

$$= m^{-1} \sum_{j=1}^m \int_1^{\exp(E_{j,n} - E_{m+1,n})} \left(\frac{a_0(e^{E_{m+1,n}}s)}{a(e^{E_{m+1,n}})} - \frac{\alpha - 1}{\alpha} s^{\alpha^{-1}}\right) \frac{ds}{s} \tag{5.5}$$

$$+ (\alpha - 1)m^{-1} \sum_{j=1}^m \left(e^{\alpha^{-1}(E_{j,n} - E_{m+1,n})} - 1\right).$$

The second term goes to 1 a.s. by Lemma 5.2. On the other hand for a given $\varepsilon > 0$ with $\alpha^{-1} + \varepsilon < 1$ and n sufficiently large, the modulus of the first term is, using Lemma 4.1, bounded above by

$$m^{-1} \sum_{j=1}^m \int_1^{\exp(E_{j,n} - E_{m+1,n})} \left(\left(\frac{\alpha - 1}{\alpha} + \varepsilon\right) s^{\alpha^{-1} + \varepsilon} - \frac{\alpha - 1}{\alpha} s^{\alpha^{-1}}\right) \frac{ds}{s}$$

$$= m^{-1} \sum_{j=1}^m \left(\frac{\alpha - 1}{\alpha} + \varepsilon\right)\left(\frac{\alpha}{1 + \alpha\varepsilon}\right)\left(e^{(\alpha^{-1} + \varepsilon)(E_{j,n} - E_{m+1,n})} - 1\right)$$

$$- m^{-1} \sum_{j=1}^m (\alpha - 1)\left(e^{\alpha^{-1}(E_{j,n} - E_{m+1,n})} - 1\right)$$

which by Lemma 5.2 converges almost surely to

$$\left(\frac{\alpha - 1}{\alpha} + \varepsilon\right)\left(\frac{\alpha}{1 + \alpha\varepsilon}\right)\left(\frac{\alpha}{1 + \alpha\varepsilon} - 1\right)^{-1} - 1 = \frac{\alpha - 1 + \alpha\varepsilon}{\alpha - 1 - \alpha\varepsilon} - 1.$$

Since $\varepsilon > 0$ is arbitrary, the first term in (5.5) goes to zero almost surely so that

$\hat{a}(n/m)a(\exp(E_{m+1,n})) \to 1$ a.s. Finally from (4.9) and Lemma 5.1, we have

$$\frac{a(\exp(E_{m+1,n}))}{a(n/m)} = \exp\left(\int_1^{\exp(E_{m+1,n}-\log(n/m))} \frac{u((n/m)x)}{x}\, dx\right) \to 1 \quad \text{a.s.}$$

which combined with $\hat{a}(n/m)/a(\exp(E_{m+1,n})) \to 1$ a.s. proves (ii).

(iii) This follows easily from (i) and (ii) and the fact that $a(n/m)/b(n/m) \to 1/(\alpha - 1)$.

(iv) Since $a^*(n/m) \to 1/\alpha$, this is immediate. □

REFERENCES

BALKEMA, A. A. and DE HAAN, L. (1972). On R. Von Mises' condition for the domain of attraction of $\exp\{-e^{-x}\}$. *Ann. Math. Statist.* **43** 1352–1354.

BILLINGSLEY, P. (1979). *Probability and Measure.* Wiley, New York.

BOOS, D. (1984). Using extreme value theory to estimate large percentiles. *Technometrics* **26** 33–39.

BREIMAN, L., STONE, C., and GINNS, J. (1979). New methods for estimating tail probabilities and extreme value distributions. Tech. Rept. #TSC-PD-A266-1, Technology Service Corp., Santa Monica, California 90405.

BREIMAN, L., STONE, C., and GINNS, J. (1981). Further developments of new methods for estimating tail probabilities and extreme value distributions. Tech Rept. #TSD-PD-A243-1, Technology Service Corp., Santa Monica, California 90405.

DUMOUCHEL, W. (1983). Estimating the stable index α in order to measure tail thickness. *Ann. Statist.* **11** 1019–1036.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications,* vol. II. Wiley, New York.

HAAN, L. DE. (1970). *On Regular Variation and Its Applications to the Weak Convergence of Sample Extremes.* Tract 32, Mathematics Center, Amsterdam.

HAEUSLER, E. and SCHNEEMEIRER, W. (1983). On parameter estimation in Pareto-type distributions. Preprint, Mathematical Institute, University of Munich.

HALL, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc. B* **44** 37–42.

HALL, W. J. and WELLNER, J. A. (1979). The rate of convergene in law of the maximum of an exponential sample. *Statistica Neerlandica* **33** 151–154.

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

MASON, D. (1981). Bounds for weighted empirical distribution functions. *Ann. Probab.* **9** 881–885.

MASON, D. (1982). Laws of large numbers for sums of extreme values. *Ann. Probab.* **10** 754–764.

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

SERFLING, R. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

SMIRNOV, N. V. (1952). Limit distributions for the terms of a variational series. *American Math. Soc. Translation Series,* Vol. II, #67.

TEUGELS, J. (1981). Limit theorems on order statistics. *Ann. Probab.* **9** 868–880.

WEISSMAN, I. (1978). Estimation of parameters and quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73** 812–815.

COLORADO STATE UNIVERSITY
DEPARTMENT OF STATISTICS
FORT COLLINS, COLORADO 80523