The Distribution of Test Statistics for Outlier Detection in Heavy-tailed Samples*

Stefan Mittnik

Institute of Statistics and Econometrics

University of Kiel

Olshausenstr. 40, D-24098 Kiel, Germany

Svetlozar T. Rachev

Institute of Statistics and Mathematical Economics

University of Karlsruhe

Kollegium am Schloss Bau II, D-76128 Karlsruhe, Germany

Gennady Samorodnitsky

School of Operations Research and Industrial Engineering

Cornell University

Rhodes Hall, Ithaca, NY 14853-3801, U.S.A.

August 12, 1999

*Parts of the work were conducted while S. Rachev visited the University of Kiel with support from the Alexander-von-Humboldt Foundation, and other parts of the work were conducted with support of the University of Karlsruhe during the visit of G. Samorodnitsky there. The research of the S. Mittnik was supported by the Deutsche Forschungsgemeinschaft (DFG) and that of of G. Samorodnitsky by NSF grants DMS-97-04982 and DMI-97-13549 and by the NSA grant MDA904-98-1-0041 at Cornell University.

Abstract

We investigate the asymptotic behavior of test statistics outliers for sample drawn from heavy–tailed distributions. We extend classical results of David, Hartley and Pearson (1954) and Grubbs (1969), who considered outlier test statistics for the finite–variance case, to the heavy–tailed infinite variance case. Our main result concerns the limiting distribution of $n^{-1/2}O_n$ for the outlier statistic

$$O_n = \frac{\max_{1\le i\le n} X_i - \min_{1\le i\le n} X_i}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}}$$

when the observations X_i are the domain of attraction of an α–stable law. We present approximate critical values for O_n for finite samples using response surface methods.

1 Introduction

Whether or not an observation in a given sample should be viewed as an outlier depends on the underlying distributional model. Although there are many practical situations where the assumption of normality is not appropriate, most of the literature on outlier detection assumes that samples are drawn from a normal distribution. Exceptions are, for example, Basu (1965), who considers the outliers problem in class of exponential distributions, Neyman and Scott (1971), who introduced the concept of outlier–proneness or outlier–resistance of families of distributions, and Green (1976), who extended these concepts to individual distributions by classifying distributions according to their outlier properties. Especially in financial applications, the normal hypothesis is frequently rejected. Already Mandelbrot (1963) observed that returns on financial assets are typically leptokurtic and fat–tailed. He suggested the stable Paretian distribution—a generalization of the normal distribution which can capture such phenomena—as a model for financial returns.[1]

In this paper, we consider the problem of outlier detection when samples are assumed to be drawn from stable Paretian distributions. More generally, we assume that samples are in the domain of attraction (DA) of a stable law. Specifically, we provide limiting distributions of the standardized maximum (or minimum) statistic of Grubbs (1969), the standardized range statistic of David, Hartley and Pearson (1954), and the standardized absolute maximum statistic. We derive critical values for the range statistic when we assume that samples stem from non–Gaussian stable Paretian distributions. To facilitate the practical implementation of outlier tests based on the range statistic, we report compact response surface approximations for finite samples which are derived from Monte Carlo simulations.

The paper organized is as follows: Section 2 gives a brief summary of stable laws. Limiting distributions of the outlier test statistics are presented in Section 3. In Section 4 we report finite–sample critical values for the standardized range statistic. Section 5 concludes.

[1] See Fama (1965), Mittnik and Rachev (1993), McCulloch (1996), Rachev and Mittnik (2000), and the references therein for discussions on this issue.

2 Stable Laws and Their Domains of Attraction

There are several ways of defining an α–stable distribution (see Zolotarev, 1986; Samorodnitsky and Taqqu, 1994, and the references therein). The classical definition, given in Lévy (1937), states that a random variable (r.v.) X is *stable*, if for any positive numbers A and B there is a positive number, C, and a real number, D, such that $AX_1 + BX_2 \stackrel{d}{=} CX + D$, where X_1 and X_2 are independent r.v.'s with $X_i \stackrel{d}{=} X, i = 1, 2$ and "$\stackrel{d}{=}$" denotes equality in distribution. For any stable r.v. X there is a number $\alpha \in (0, 2]$ such that C satisfies $C^\alpha = A^\alpha + B^\alpha$ (see Feller, 1971, Sec. 17.4). The exponent α is called the *index of stability*. For $\alpha < 2$ a non–degenerate stable r.v. X with index of stability α satisfies $P(|X| > t) \sim ct^{-\alpha}$ for some $c > 0$ as $t \to \infty$, and the left and right tails of X are balanced as in (4) below. Hence, if $\alpha < 2$, the tails of the distribution of a stable r.v. are fatter than those of the normal distribution; and the tail–thickness increases as α decreases. This is why α is also referred to as the *tail–thickness parameter*. If $\alpha < 2$, moments of order α or higher do not exist. A stable r.v. with index α is said to be α–stable. A Gaussian random variable is a 2–stable random variable (i.e., $\alpha = 2$). Indeed, if X_1 and X_2 are independent normal with a common mean μ and variance σ^2, then $AX_1 + BX_2 \sim N((A + B)\mu, (A^2 + B^2)\sigma^2)$; i.e., we have $C = (A^2 + B^2)^{\frac{1}{2}}$ and $D = (A + B - C)\mu$.

Closed–form expressions of α–stable distributions or their densities exist only in few special cases. However, the logarithm of the characteristic function (ch.f.), $f(\theta) = Ee^{i\theta X}$, of α–stable r.v. X, can be written as

$$\ln f(\theta) = \begin{cases} -\sigma^\alpha |\theta|^\alpha [1 - i\beta \,\mathrm{sign}(\theta) \tan\frac{\pi\alpha}{2}] + i\mu\theta, & \text{for } \alpha \neq 1, \\ -\sigma|\theta|[1 + i\beta\frac{\pi}{2} \,\mathrm{sign}(\theta) \ln|\theta|] + i\mu\theta, & \text{for } \alpha = 1, \end{cases} \tag{1}$$

$\theta \in \mathbf{R}$, where $\mu \in \mathbf{R}$ is the *location parameter*; $\sigma \geq 0$ is the *scale parameter*; and $\beta \in [-1, 1]$ is the *skewness parameter*. The distribution function of an α–stable r.v. satisfying (1) is denoted by $S(x; \alpha, \beta, \sigma, \mu)$. If $\beta = 0$, the distribution is symmetric. The location parameter shifts the distribution to the left or right, while for $\alpha \neq 1$ or $\beta = 0$ the scale parameter expands or contracts it about μ. If X has ch.f. (1) we write $X \stackrel{d}{=} S_\alpha(\beta, \sigma, \mu)$. For $\alpha = 2$, $S_2(\beta, \sigma, \mu)$ is the normal distribution $N(\mu, 2\sigma^2)$.

Unless both $\alpha = 1$ and $\beta \neq 0$ the standardized version $(X-\mu)/\sigma$ of $X \stackrel{d}{=} S_\alpha(\beta, \sigma, \mu)$ has distribution $S_\alpha(\beta, 1, 0)$.

A sample $U_1, U_2, \ldots$ of i.i.d. observations is said to be *in the domain of attraction of an α–stable law* with index $\alpha \in (0,2]$ if there exist constants $a_n \geq 0$ and $b_n \in \mathbf{R}$ such that

$$a_n^{-1} S_n - b_n \stackrel{w}{\Rightarrow} X, \tag{2}$$

where $S_n = U_1 + \cdots + U_n$, X is a non–degenerate α–stable r.v., and "$\stackrel{w}{\Rightarrow}$" stands for weak convergence. In particular, when U_i's are α–stable, $U_1 \stackrel{d}{=} S_\alpha(\beta, \sigma, \mu)$, (2) holds and, moreover, we have $a_n^{-1} S_n - b_n \stackrel{d}{=} U_1$, with $a_n = n^{1/\alpha}$ and $b_n = \mu(n^{1-1/\alpha} - 1)$ for $\alpha \neq 1$, and $b_n = \frac{2}{\pi}\sigma\beta n \ln n$ for $\alpha = 1$.

The assumption that the disturbances U_i's are in the domain of attraction of an α–stable law is, hence, a relaxation of the assumption of α–stable distributed disturbances. In fact, for $\alpha < 2$ the domain–of–attraction condition (2) is equivalent to the assumption that the tail behavior of U_i is of the Pareto–Lévy form (cf. Feller, 1971, p. 303):

$$P(|U_i| > t) = t^{-\alpha} L(t), \quad t > 0, \tag{3}$$

where $L(t)$ is a slowly varying function as $t \to \infty$,[2] and

$$\lim_{t\to\infty} \frac{P(U_i > t)}{P(|U_i| > t)} = p, \quad \lim_{t\to\infty} \frac{P(U_i < -t)}{P(|U_i| > t)} = q, \tag{4}$$

for some $p \geq 0$ and $q \geq 0$ with $p + q = 1$.

We shall further assume that U_i are in the *domain of normal attraction of an α–stable law*, that is, for some $c > 0$,

$$P(|U_i| > t) \sim ct^{-\alpha} \quad \text{as } t \to \infty,$$

and furthermore the limiting relationships (4) hold.[3]

[2] $L(t)$ is a slowly varying function as $t \to \infty$, if for every constant $c > 0$, $\lim_{t\to\infty} L(ct)/L(t)$ exists and is equal to 1. The definition is similar for a function of a discrete variable $l(n)$, $n = 1, 2, \ldots$. We will use L or l to denote a slowly varying function.

[3] The U_i's are in the domain of normal attraction of an α–stable law, if (2) holds with $a_n = c_0 n^{1/\alpha}$ for some positive constant c_0. Note that when the U_i' are in the *general* domain of attraction, then, in (2), $a_n = n^{1/\alpha} l(n)$ for some slowly varying function $l(n)$ as $n \to \infty$.

3 Asymptotic Results for Tests of Outlier

We now consider the asymptotic distributions of statistics for outlier detection. Grubbs (1969) proposed the standardized maximum,

$$M_n = \frac{\max_{1\le i\le n} X_i - \bar{X}}{S_x}, \tag{5}$$

and the standardized minimum,

$$m_n = \frac{\bar{X} - \min_{1\le i\le n} X_i}{S_x}; \tag{6}$$

of a sample as a test statistics, where, $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$ is the sample mean and $S_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$ is the consistent sample standard deviation. A modification of (5) and (6) is given by

$$\bar{M}_n = \frac{\max_{1\le i\le n}|X_i| - \bar{X}}{S_x} = \max(M_n, -m_n). \tag{7}$$

The next two propositions underline the drastic change in the asymptotic behavior of statistics (5)-(7) when the normality assumption is replaced by the non-normal stable assumption (see also Rachev, Mittnik and Kim (1996)).

Proposition 1 (The normal case) *Suppose $\{X_i\}_{i\ge 1}$ are i.i.d. standard normals. Then*

$$\sqrt{2\log n}M_n - B_n \xrightarrow{d} \Lambda, \tag{8}$$

$$\sqrt{2\log n}m_n - B_n \xrightarrow{d} \Lambda, \tag{9}$$

and

$$\sqrt{2\log n}\bar{M}_n - B_n - \log 2 \xrightarrow{d} \Lambda, \tag{10}$$

where $B_n = 2\log n - \frac{1}{2}\log\log n - \frac{1}{2}\log 4\pi$ and Λ has the Gumbel extreme-value distribution, i.e.

$$F_\Lambda(x) = P(\Lambda < x) = \exp\left\{-e^{-x}\right\},\ -\infty < x < \infty.$$

Note that for the non-Gaussian stable case with $0 < \alpha \le 1$ we cannot center with $\bar{X}$, because $\bar{X}$ becomes unstable as $n \to \infty$. Then, one would have to replace $\bar{X}$ by the empirical median or trimmed mean when defining statistics M_n, m_n and $\bar{M}_n$. Since the infinite-mean-case $0 < \alpha \le 1$ has little practical importance, we assume in the following that $\alpha > 1$.

Proposition 2 (The stable case) *Suppose that* $(X_i)_{i\geq 1}$ *are in the domain of attraction of an* α*-stable law with* $1<\alpha<2$*. Then, for any* $\varepsilon>0$*,*

$$n^{-\frac{1}{2}+\varepsilon}M_n \xrightarrow{P} \infty, \quad n^{-\frac{1}{2}+\varepsilon}m_n \xrightarrow{P} \infty, \quad n^{-\frac{1}{2}+\varepsilon}\bar{M}_n \xrightarrow{P} \infty$$

and

$$n^{-\frac{1}{2}-\varepsilon}M_n \xrightarrow{P} 0, \quad n^{-\frac{1}{2}-\varepsilon}m_n \xrightarrow{P} 0, \quad n^{-\frac{1}{2}-\varepsilon}\bar{M}_n \xrightarrow{P} 0.$$

Propositions 1 and 2 show that the proper normalizations for M_n, m_n and $\bar{M}_n$ are, in fact, quite different for the normal and the non–Gaussian stable case.

Our main result in this section concerns the exact limiting distribution and refines the assertion in Proposition 2. Further, we shall study the asymptotic behavior of the symmetrized outlier test statistic

$$O_n = M_n + m_n = \frac{\max_{1\leq i\leq n} X_i - \min_{1\leq i\leq n} X_i}{S_x}. \tag{11}$$

Suppose that the observations $(X_i)_{i\geq 1}$ are i.i.d. random variables in the domain of attraction of the α-stable law with index $\alpha\in(1,2)$, that is, relations (2) – (4) hold. Given the tail masses p and q defined by (4), consider a random variable N with a geometric–type distribution:

$$P(N=k) = pq^{k-1} + qp^{k-1},\ k\geq 2. \tag{12}$$

In particular, if $p=0$ or 1, then $N=+\infty$, a.s. and so $\Gamma_N^{-1/\alpha}=0$ a.s.. Next let $(\Gamma_i)_{i\geq 1}$ be the sequence of arrivals of a standard Poisson process, that is,

$$\Gamma_i = e_1 + \ldots + e_i,$$

where $(e_j)_{j\geq 1}$, are i.i.d. exponential random variables with mean 1. We assume that $(\Gamma_i)_{i\geq 1}$ is independent of N and introduce the random variable

$$U^* := \frac{\Gamma_1^{-1/\alpha} + \Gamma_N^{-1/\alpha}}{(\sum_{i=1}^{\infty}\Gamma_i^{-2/\alpha})^{1/2}} \tag{13}$$

We can now formulate our main result.

Theorem 1 *Suppose that* $(X_i)_{i\geq 1}$ *are in the domain of normal attraction of the* α*-stable law with* $1<\alpha<2$*. Then, as* $n\to\infty$*,*

$$n^{-1/2}O_n \overset{w}{\Rightarrow} U^*. \tag{14}$$

Remark 1 *An alternative representation of the limiting distribution for statistic* O_n *is given by*

$$U^* \stackrel{d}{=} \frac{p^{1/\alpha}(\Gamma_1^{(1)})^{-1/\alpha} + q^{1/\alpha}(\Gamma_1^{(2)})^{-1/\alpha}}{\left(p^{2/\alpha}\sum_{i=1}^{\infty}(\Gamma_i^{(1)})^{-2/\alpha} + q^{2/\alpha}\sum_{i=1}^{\infty}(\Gamma_i^{(2)})^{-2/\alpha}\right)^{1/2}}, \tag{15}$$

where $(\Gamma_i^{(1)})_{i\geq 1}$ *and* $(\Gamma_i^{(2)})_{i\geq 1}$ *are two independent sequences of standard Poisson arrivals. Note also that the denominator in the two representations of the limiting law* U^* *in (13) and (15) is the square root of a positive* $\alpha/2$*–stable random variable.*

Remark 2 *An argument identical to the one used in the proof of Theorem 1 below gives limiting distributions of the statistics* M_n*,* m_n *and* $\bar{M}_n$*. In fact,*

$$n^{-1/2}M_n \stackrel{w}{\Rightarrow} U_+^*, \tag{16}$$

with

$$U_+^* := \frac{\Gamma_{N_1}^{-1/\alpha}}{(\sum_{i=1}^{\infty}\Gamma_i^{-2/\alpha})^{1/2}}, \tag{17}$$

where N_1 *is a geometric random variable*

$$P(N_1 = k) = p(1-p)^{k-1},\ k = 1, 2, \ldots,$$

independent of the Poisson process $(\Gamma_i)_{i\geq 1}$*,*

$$n^{-1/2}m_n \stackrel{w}{\Rightarrow} U_-^*, \tag{18}$$

with

$$U_-^* := \frac{\Gamma_{N_2}^{-1/\alpha}}{(\sum_{i=1}^{\infty}\Gamma_i^{-2/\alpha})^{1/2}}, \tag{19}$$

where N_2 *is a geometric random variable*

$$P(N_2 = k) = q(1-q)^{k-1},\ k = 1, 2, \ldots,$$

independent of the Poisson process $(\Gamma_i)_{i\geq 1}$*, and*

$$n^{-1/2}\bar{M}_n \stackrel{w}{\Rightarrow} U_\pm^*, \tag{20}$$

with

$$U_\pm^* := \frac{\Gamma_1^{-1/\alpha}}{(\sum_{i=1}^{\infty}\Gamma_i^{-2/\alpha})^{1/2}}. \tag{21}$$

Furthermore, alternative representations of the type (15) are available for both U_+^* *and* U_-^**.*

Proof of Theorem 1. Let $g(x) = 1/P(|X_i| > x)$ and consider the generalized inverse of $g(x)$:

$$g^{\leftarrow}(y) = sup\{x : g(x) \leq y\}.$$

Set $a_n = g^{\leftarrow}(n),\ n \geq 1$, then, as $n \to \infty$, $a_n \sim cn^{-1/\alpha}$.[4]

Before continuing, we need some basic definitions and results on Poisson random measures, in short, PRM, (see Resnick, 1987). Let E be a locally compact topological space with a countable base; E plays the role of the state space for the point processes under consideration. Let $\mathcal{E}$ be the Borel σ-algebra of subsets of E. A *point measure, m,* on $\mathcal{E}$ with support $\{x_i, i \geq 1\} \subset E$ is defined by

$$m = \sum_{i=1}^{\infty} \epsilon_{x_i},$$

where, for any $x \in E$, ϵ_x is the unit mass in x:

$$\epsilon_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}, \ A \in \mathcal{E}\,.$$

A *point process* N on E is a random element, $N : (\Omega, \mathcal{A}, P) \to (M_P(E), \mathcal{M}_p(E))$, on the original probability space $(\Omega, \mathcal{A}, P)$ with values in the space $M_P(E)$ of all point measures on E with the σ- algebra $\mathcal{M}_P(E)$ generated by the sets $\{m \in M_P(E) : m(F) \in B\}$, $F \in \mathcal{E}$, and B is a Borel set in $[0, \infty]$, i.e., B $\in \mathcal{B}([0, \infty])$.

Let μ be a Radon measure on $(E, \mathcal{E})$, that is, μ is finite on all compact subsets of E. A point process N is called *Poisson random measure (PRM)* with mean measure μ if,

(i) for every $E \in \mathcal{E}$, and every $k = 0, 1, 2, \ldots$,

$$P(N(E) = k) = \frac{\mu(E)^k}{k!} e^{-\mu(E)},$$

if $\mu(E) < \infty$, and $P(N(E) = k) = 0$, if $\mu(E) = \infty$; and

(ii) if $E_1, \ldots, E_k$ (for $k = 1, 2, \ldots$) are mutually disjoint finite measure sets in $\mathcal{E}$, then N$(E_1), \ldots,$ N(E_k) are independent random variables.

[4] Here, and in what follows, c stands for a generic positive constant which can be different in different contexts.

Next, consider an array of random variables $(X_{n,j}, j \geq 1, n \geq 1)$ with values in $(E, \mathcal{E})$ and assume that, for each n, $(X_{n,j})_{j\geq 1}$ are i.i.d. Suppose that the Radon measure

$$\mu_n(A) := nP(X_{n,1} \in A),\ A \in \mathcal{E},$$

converges vaguely to a Radon measure μ on $(E, \mathcal{E})$. Recall that $(\mu_n)_{n\geq 1}$ converges vaguely to μ ($\mu_n \stackrel{v}{\to} \mu$), if $\int f\, d\mu_n$ converges as $n \to \infty$ to $\int f\, d\mu$ for every nonnegative continuous function f with a compact support.

Lemma 1 *(See Resnick, 1987, Proposition 3.21). Let* $\xi_n = \sum_{k\geq 1} \varepsilon_{(\frac{k}{n}, X_{k,n})}$ *and* ξ *be a PRM on* $[0, \infty) \times E$ *with mean measure* $dt \times d\mu$. *Then,*

$$\mu_n \stackrel{v}{\to} \mu \tag{22}$$

implies the weak convergence

$$\xi_n \stackrel{w}{\Rightarrow} \xi. \tag{23}$$

Remark 3 *In (23),* $\stackrel{w}{\Rightarrow}$ *stands for the weak convergence of stochastic point processes, that is, the weak convergence in the space* $M_P([0, \infty) \times E)$.

We now apply the above lemma to a sequence $(X_i)_{i\geq 1}$ of i.i.d. random variables in the domain of attraction of an α-stable law. Namely, we set, in Lemma 1, $X_{k,n} = \frac{X_k}{a_n}$, where a_n was defined as $g^{\leftarrow}(n)$. From this point on for the remainder of this paper we use $E = [-\infty, \infty] \setminus \{0\}$, unless explicitly mentioned otherwise. Then, as $n \to \infty$,

$$\sum_{k=1}^{\infty} \varepsilon_{(\frac{k}{n}, \frac{X_k}{a_n})} \stackrel{w}{\Rightarrow} X^* = \sum_i \varepsilon_{(t_i, j_i)} \qquad \text{in } M_P([0, \infty) \times E), \tag{24}$$

where the limit is a PRM with mean measure $dt \times d\nu$, and

$$\nu(dx) = \alpha p x^{-(1+\alpha)} dx \mathbf{1}_{\{x>0\}} + \alpha q |x|^{-(1+\alpha)} dx \mathbf{1}_{\{x>0\}} \tag{25}$$

(see formula (4.70) in Resnick, 1987, p. 226).

Observe that, for $\mathbf{X} \in M_P([0, \infty] \times E)$,

$$M_+(\mathbf{X}) = \max\{j : (t, j) \in \mathbf{X},\ t \leq 1\} \tag{26}$$

is a functional that is continuous except on a set with measure 0 with respect to the law of the PRM in the right–hand–side of (24). A similar argument is valid for

$$M_-(\mathbf{X}) = \min\{j : (t,j) \in \mathbf{X}, t \le 1\}. \tag{27}$$

The continuity of M_+ is stated and proved in Resnick (1987, pp. 211, 214). The situation with respect to M_- is analogous. The above continuity is established in Resnick (1987) for M_+ when $p > 0$, and, thus, for M_- when $q > 0$. This leaves only the case of, say, M_- with $q = 0$ to be considered. In this case, we interpret M_- as $M_-(m) = 0$ for every $m \in M_P([0,\infty] \times E)$ such that $m([0,\infty) \times [-\infty, 0)) = 0$ and $m((0,1) \times (0,\varepsilon]) = \infty$ for all $\varepsilon > 0$. Observe that, for $q = 0$ the PRM $\mathbf{X}^*$ in (24) has a.s. the latter property. Furthermore, if m has this property and $m_n \in M_P([0,\infty) \times E)$, $n \ge 1$ are such that $m_n \xrightarrow{v} m$, then, m_n will eventually have points arbitrarily close to 0, and no points below a given positive distance from 0, implying that $M_-(m_n) \overset{w}{\Rightarrow} M_-(m)$. Therefore, M_+ and M_- are a.s. continuous.

Furthermore, for $\mathbf{X} \in M_P([0,\infty] \times E)$ and $0 < a < b < \infty$, let

$$I_{a,b}^{(2)}(\mathbf{X}) = \int_{t \le 1} j^2 \mathbf{1}(a \le |j| \le b) d\mathbf{X}. \tag{28}$$

It follows from Resnick (1987, Exercise 4.4.2.8 (c)) that $I_{a,b}^{(2)}$ is a map

$$I_{a,b}^{(2)}(\mathbf{X}) : M_P([0,\infty) \times E) \to \mathbf{R}_+,$$

that is, a.s. continuous with respect to the law of $\mathbf{X}^*$. Therefore, for every $0 < a < b < \infty$,

$$\frac{M_+(\mathbf{X}) - M_-(\mathbf{X})}{(I_{a,b}^{(2)}(\mathbf{X}))^{1/2}} \tag{29}$$

is a.s. continuous, $M_P([0,\infty) \times E) \to \mathbf{R}$, and so by the Continuous Mapping Theorem, as $n \to \infty$, relationship (24) implies

$$\frac{M_+(X_n) - M_-(X_n)}{(I_{a,b}^{(2)}(X_n))^{1/2}} \overset{w}{\Rightarrow} \frac{M_+(\mathbf{X}^*) - M_-(\mathbf{X}^*)}{(I_{a,b}^{(2)}(\mathbf{X}^*))^{1/2}}, \tag{30}$$

where $X_n = \sum_{k \ge 1} \varepsilon_{(\frac{k}{n}, \frac{X_k}{a_n})}$. Now,

$$M_+(\mathbf{X}_n) = \frac{1}{a_n} \max_{i \le n} X_i, \quad M_-(\mathbf{X}_n) = \frac{1}{a_n} \min_{i \le n} X_i, \tag{31}$$

and

$$I^{(2)}_{a,b}(\mathbf{X}_n) = \frac{1}{a_n^2}\sum_{i=1}^n X_i^2 \mathbf{1}(aa_n \le |X_i| \le ba_n). \tag{32}$$

Furthermore, the points of $\mathbf{X}^*$ on $\{t \le 1\}$ arranged in the non–increasing order by the magnitude of the "jumps" j_i's can be represented in distribution as $(U_j^o, \delta_j \Gamma_j^{-1/\alpha})$, $j \ge 1$, where $(U_j^o, j \ge 1), (\delta_j, j \ge 1)$ and $(\Gamma_j, j \ge 1)$ are three independent sequences of random variables; $(U_j^0, j \ge 1)$ are i.i.d. uniformly distributed on $[0,1]$, i.e., $U_j^0 \stackrel{d}{=} U(0,1), (\delta_j, j \ge 1)$ are i.i.d. Bernoulli r.v.'s, $P(\delta_j = 1) = 1 - P(\delta_j = -1) = p$, and finally $(\Gamma_j, j > 1)$ are the standard Poisson arrivals. Therefore, by (31),

$$\frac{\max_{i\le n} X_i - \min_{i \le n} X_i}{\left(\sum_{i=1}^n X_i^2 \mathbf{1}(aa_n \le |X_i| \le ba_n)\right)^{1/2}} \stackrel{w}{\Rightarrow} \frac{\Gamma_1^{-1/\alpha} + \Gamma_N^{-1/\alpha}}{\left(\sum_{i=1}^\infty \Gamma_i^{-2/\alpha} \mathbf{1}(a \le \Gamma_i^{-1/\alpha} \le b)\right)^{1/2}} := U^*(a,b), \tag{33}$$

where N is independent of $(\Gamma_j, j \ge 1)$ and has the law given in (12). An alternative representation of $U^*(a,b)$ is obtained by splitting the marked Poisson process $(\delta_j, \Gamma_j), j \ge 1$ into two independent parts, that with $\delta_j = 1$ and that with $\delta_j = -1$. This leads to a representation

$$U^*(a,b) = \left(p^{1/\alpha}(\Gamma_1^{(1)})^{-1/\alpha} + (1-p)^{1/\alpha}(\Gamma_1^{(2)})^{-1/\alpha}\right) \tag{34}$$

$$\left(p^{2/\alpha}\sum_{i=1}^\infty (\Gamma_i^{(1)})^{-2/\alpha}\mathbf{1}(a \le (\Gamma_i^{(1)})^{-1/\alpha} \le b)\right.$$

$$\left. + (1-p)^{2/\alpha}\sum_{i=1}^\infty (\Gamma_i^{(2)})^{-2/\alpha}\mathbf{1}(a \le (\Gamma_i^{(2)})^{-1/\alpha} \le b)\right)^{-1/2},$$

where $(\Gamma_j^{(1)}\ j \ge 1)$ and $(\Gamma_j^{(2)}\ j \ge 1)$ are independent sequences of standard Poisson arrivals.

Our goal is to prove that

$$\frac{\max_{i\le n} X_i - \min_{i \le n} X_i}{\left(\sum_{i=1}^n X_i^2\right)^{1/2}} \stackrel{w}{\Rightarrow} \frac{\Gamma_1^{-1/\alpha} + \Gamma_N^{-1/\alpha}}{\left(\sum_{i=1}^\infty \Gamma_i^{-2/\alpha}\right)^{1/2}} := U^*. \tag{35}$$

To this end, it is enough to show that

$$\frac{\sum_{i=1}^n X_i^2}{(\max_{i\le n} X_i - \min_{i \le n} X_i)^2} \stackrel{w}{\Rightarrow} (U^*)^{-2}. \tag{36}$$

It follows from (33) that, for $0 < a < 1$,

$$\frac{\sum_{i=1}^n X_i^2 \mathbf{1}\left(aa_n \le |X_i| \le a^{-1}a_n\right)}{(\max_{i\le n} X_i - \min_{i\le n} X_i)^2} \overset{w}{\Rightarrow} (U^*(a,a^{-1}))^{-2} \tag{37}$$

and that $U^*(a,a^{-1}) \overset{w}{\Rightarrow} U^*$, as $a \to 0$. Therefore, limiting relationship (36) will follow from the continuity Theorem 4.2 of Billingsley (1968), once we prove that

$$\lim_{a\to 0} \limsup_{n\to\infty} \; P(|\triangle_n(a)| \ge \varepsilon) = 0 \tag{38}$$

for all $\varepsilon \in (0,1)$, where

$$\triangle_n(a) = \frac{\sum_{i=1}^n X_i^2}{(\max_{\,i\le n} X_i - \min_{\,i\le n} X_i)^2} - \frac{\sum_{i=1}^n X_i^2 \mathbf{1}(aa_n \le |X_i| \le a^{-1}a_n)}{(\max_{\,i\le n} X_i - \min_{\,i\le n} X_i)^2}. \tag{39}$$

We have

$$\begin{aligned} P(|\triangle_n(a)| \ge \varepsilon) \quad &\le P\left(\left(\frac{1}{a_n}\left(\max_{i\le n} X_i - \min_{i\le n} X_i\right)\right)^{-1} \ge \varepsilon^{-1/4}\right) \\ &+P\left(\frac{1}{a_n^2}\sum_{i=1}^n X_i^2\mathbf{1}(|X_i| < aa_n) \ge \frac{1}{2}\varepsilon^{3/2}\right) \\ &+P\left(\frac{1}{a_n^2}\sum_{i=1}^n X_i^2\mathbf{1}\left(|X_i| > a^{-1}a_n\right) \ge \frac{1}{2}\varepsilon^{3/2}\right) \\ &:= P_1^{(n)}(\varepsilon) + P_2^{(n)}(a,\varepsilon) + P_3^{(n)}(a,\varepsilon) \end{aligned} \tag{40}$$

Now,

$$\frac{1}{a_n}(\max_{i\le n} X_i - \min_{i\le n} X_i) \text{ converges weakly to an a.s. positive random variable.} \tag{41}$$

Therefore,

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} P_1^{(n)}(\varepsilon) = 0. \tag{42}$$

Furthermore,

$$\begin{aligned} E(a_n^{-2})\sum_{i=1}^n X_i^2\mathbf{1}(|X_i| < aa_n) \quad &\le cn^{-2/\alpha} nE(X_1^2\mathbf{1}(|X_1| < an^{1/\alpha})) \\ &= cn^{-2/\alpha+1}\int_0^{an^{1/\alpha}} x^2 F_{|X_1|}(dx) \\ &= cn^{-2/\alpha+1}\int_0^{an^{1/\alpha}} 2yP(y < |X_i| \le an^{1/\alpha})\,dy \\ &\le 2cn^{-2/\alpha+1}\int_0^{an^{1/\alpha}} P(|X_1| > y)y\,dy \\ &\le cn^{-2/\alpha+1}\int_0^{an^{1/\alpha}} y^{-\alpha+1}\,dy = ca^{2-\alpha}. \end{aligned}$$

Therefore,

$$P_2^{(n)}(a,\varepsilon) \le \frac{ca^{2-\alpha}}{\varepsilon^{3/2}}, \quad \text{for all } a \text{ and } n,\ \epsilon$$

and so

$$\lim_{a\to 0}\limsup_{n\to\infty} P_2^{(n)}(a,\varepsilon) = 0. \tag{43}$$

Furthermore, as $n \to \infty$,

$$\begin{aligned}
P_3^{(n)}(a,\varepsilon) &\le P(\text{at least one of } X_is,\ i=1,\dots,n \text{ satisfies } |X_i| > a^{-1}a_n) \\
&= 1-(P(|X_i| \le a^{-1}n^{1/\alpha}))^n \\
&= 1-(1-P(|X_i| > a^{-1}n^{1/\alpha}))^n \\
&\le 1-(1-ca^{\alpha}n^{-1})^n \\
&\to 1-e^{-ca^{\alpha}}.
\end{aligned}$$

Therefore,

$$\lim_{a\to 0}\limsup_{n\to\infty} P_3^{(n)}(a,\varepsilon) = 0. \tag{44}$$

Then, (38) follows from (42), (43) and (44). Hence, (36) and, thus, (35) have been proved.

Next,

$$\left(n^{1/2}\frac{S_x}{\max_{i\le n}X_i - \min_{i\le n}X_i}\right)^2 = \frac{(\sum_{i=1}^n X_i^2)}{(\max_{i\le n}X_i - \min_{i\le n}X_i)^2} - \frac{n(\bar{X})^2}{(\max_{i\le n}X_i - \min_{i\le n}X_i)^2}. \tag{45}$$

Since $\bar{X} \to_{n\to\infty} E(X)$ a.s. and

$$a_n^{-1}(\max_{i\le n}X_i - \min_{i\le n}X_i) \stackrel{w}{\Rightarrow} W_{\mathbf{1}}, \tag{46}$$

where W_1 is a non-degenerate a.s. positive r.v., it follows that

$$\frac{n(\bar{X})^2}{(\max_{i\le n}X_i - \min_{i\le n}X_i)^2} = (\bar{X})^2\frac{1}{(a_n^{-1}(\max_{i\le n}X_i - \min_{i\le n}X_i))^2}\frac{n}{a_n^2} \to_{n\to\infty} 0 \tag{47}$$

in probability, because $a_n \sim cn^{-1/\alpha}$ as $n\to\infty$, and $\alpha < 2$. It follows from (36) and (45) – (47) that

$$\left(n^{1/2}\frac{S_x}{\max_{i\le n}X_i - \min_{i\le n}X_i}\right)^2 \stackrel{w}{\Rightarrow} (U^*)^{-2}.$$

implying that $n^{-1/2}\frac{\max_{i\le n} X_i - \min_{i\le n} X_i}{S_x} \overset{w}{\Rightarrow} U^*$, as desired. □

It is important to compare the test–statistic O_n with its L_1–version

$$O_n^* := \frac{\max_{1\le i\le n} X_i - \min_{1\le i\le n} X_i}{\frac{1}{n}\sum_{i=1}^n |X_i|} \tag{48}$$

when $1 < \alpha < 2$. From (41) we see that

$$n^{-1/\alpha} O_n^* \overset{w}{\Rightarrow} \frac{cW_1}{E|X_1|} \tag{49}$$

for some $c > 0$. Comparing the limiting results for O_n and O_n^* (cf. (14) and (49)) we conclude that the self–normalization of O_n avoid the use of the often unknown or imprecisely known tail parameter α.

4 Simulation of Finite Sample Behavior

We simulated test statistic O_n in (11) finite samples with random variables from symmetric standard α–stable distributions with the index of stability, α, assuming values 1.01, 1.05, 1.1, 1.2, ..., 1.8, 1.9, 1.95, 1.99 for sample sizes n = 10, 20, 30, 40, 50, 75, 100, 200, 300, 400, 500, 750, 1,000, 2,000, 3,000, 4,000, 5,000, 7,500, and 10,000. For each of the 247 (α, n)–combinations we generated 30,000 replication.[5] The sample quantiles with q = .90, .95, .99, .995 are shown in Figure 1.

Rather than tabulating results for selected sample sizes and α values, we employ response surface techniques to summarize our simulation results in a compact fashion.[6] By doing so, we obtain close approximations as was to be expected given the smoothness of the simulated critical values shown in Figure 1. Specifically, for the response surfaces the functional form

$$cv_q(\alpha, n) \approx \sum_{i=0}^{2}\sum_{j=0}^{1} c_{q,ij}\bar{\alpha}^i n^{j/2}, \qquad 1.01 \le \alpha \le 1.99, \tag{50}$$

[5] The pseudo random variates for the α–stable distributions were generated with the algorithm of Weron (1996).

[6] Response surface methodology has been used in various statistical and econometric applications; see, for example, Myers et al. (1989).

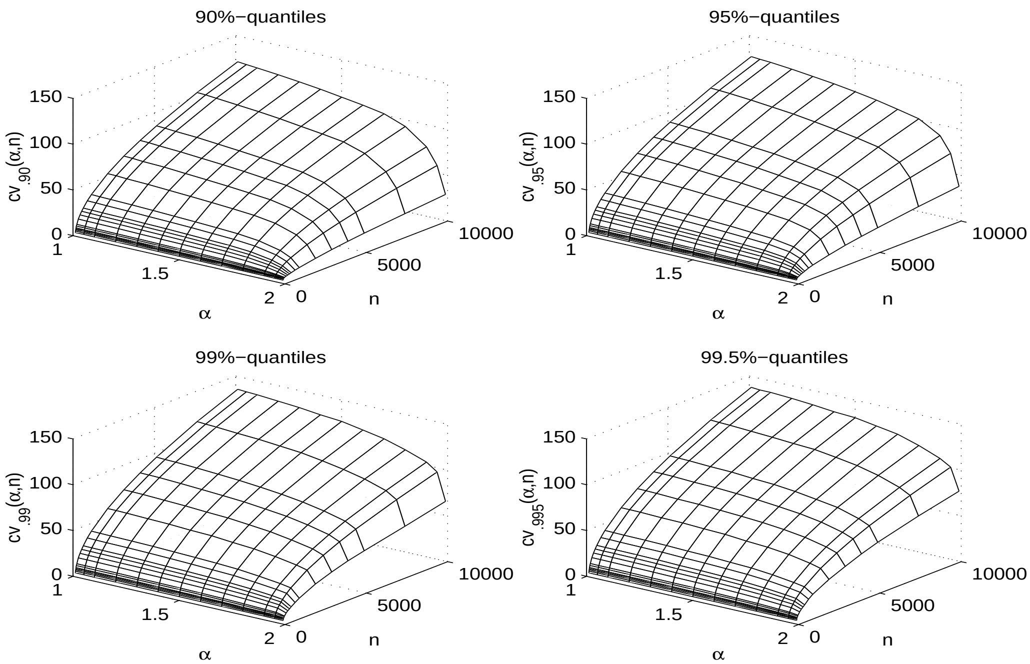

Figure 1: Simulated Critical Values for Outlier Statistic O_n

was specified, where $cv_q(\alpha, n)$ denotes the approximate quantiles for a given (α, n)–pair; and

$$\bar{\alpha} = \ln(2 - \alpha).$$

For each q–value, the estimation is based on 247 simulated data points. The estimation results for the coefficients $c_{q,ij}$ are given in Table 1. The $\bar{R}^2$–values reported in Table 1 indicate that the response surface approximations based on only six coefficients yield rather close fits to the simulated values.

We also simulated the range statistic O_n for the random variables from a standard normal distribution (i.e., $\alpha = 2$) and obtain close matches to the critical values reported in David, Hartley and Pearson (1954).[7] The critical values for $\alpha < 2$, behave quite differently from those for $\alpha = 2$, in that they increase rather slowly as n increases when $\alpha = 2$. Fitting the response surface

$$cv_q(2, n) \approx d_0 + d_1 \ln n + d_2 \ln \ln n \tag{51}$$

[7] They report all the quantiles we consider, but their sample sizes range from $n = 3$ to $n = 1,000$.

Table 1: Response Surface Estimates for $1 < \alpha < 2^a$

Coefficient	q .90	.95	.975	.99	.995
$c_{q,00}$	0.0099	-0.0002	0.0071	-0.0247	-0.0538
$c_{q,01}$	1.2252	1.2717	1.3125	1.3598	1.3875
$c_{q,10}$	0.1131	-0.0179	0.0542	0.0769	0.0805
$c_{q,11}$	0.1936	0.1574	0.1377	0.1309	0.1279
$c_{q,20}$	0.1202	0.1241	0.1200	0.0992	0.0841
$c_{q,21}$	-0.0047	-0.0098	-0.0102	-0.0057	-0.0021
$\bar{R}^2$	.99976	.99992	.99988	.99992	.99997

[a]The entries are the least squares estimates of coefficients c_{ij} in (50). $\bar{R}^2$ is the adjusted R^2–value.

to the simulated quantiles, we obtain the OLS estimates reported in Table 2. As the $\bar{R}^2$–values indicate, the fit of (51) is extremely close. If at all necessary, approximate critical values for $1.99 < \alpha < 2$ could be obtained via linear interpolation between $cv_q(1.99, n)$ and $cv_q(2, n)$.

5 Conclusions

We have considered the problem of outlier–detection when samples are assumed to be drawn from distributions in the DA of the α–stable law. Limiting distributions for certain test statistics have been derived. Approximate critical values for the standardized range statistic have been established using response surface methods. The simple functional form of the response surface facilitates practical implementation of outlier–detection procedures in heavy–tailed samples.

The results are of particular interest in financial modeling. It is well known that returns on financial assets (e.g., stock-price changes) have heavy-tailed distributions. For example, the question of whether an observation (e.g., a market crash) should be viewed as a "regular event" or as an "exceptional outlier" can be of great importance

Table 2: Response Surface Estimates for $\alpha = 2^a$

Coefficient	q .90	.95	.975	.99	.995
d_0	1.1135	1.0859	1.0381	0.9225	0.8115
d_1	0.2063	0.1507	0.0976	0.0202	-0.0380
d_2	2.3675	2.6988	3.0224	3.4937	3.8564
$\bar{R}^2$	.999998	.999999	.999999	.999997	.999996

[a]The entries are the least squares estimates of coefficients d_i in (51). $\bar{R}^2$ is the adjusted R^2–value.

in financial decision making or in empirical modeling.

References

[1] Basu, A.P. (1965), On some tests of hypotheses relating to the exponential distribution when some outliers are present. *Journal of the American Statistical Association* 60, 548–559.

[2] Billingsley, P. (1968), *Convergence of Probability Measures*, New York: Wiley & Sons.

[3] David, H.A., H.O. Hartley & E.S. Pearson (1954), The distribution of the ratio, in a single normal sample, of range to standard deviation. *Biometrika* 41, 482–493.

[4] Fama, E. (1965), The behavior of stock market prices, *Journal of Business* 38, 34–105.

[5] Feller, W. (1971), *An Introduction to Probability Theory and Its Application* Vol. 2 (2nd ed.), New York: Wiley.

[6] Green, R.F. (1976), Outlier–prone and outlier–resistant distributions. *Journal of the American Statistical Association* 71, 502–505.

[7] Grubbs, F.E. (1969), Procedures for detecting outlying observations in samples. *Technometrics* 11, 1–21.

[8] Mandelbrot, B. (1963), The variation of certain speculative prices, *Journal of Business* 26, 394-419.

[9] McCulloch, J.H. (1996), Financial applications of stable distributions. G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics*, Vol. 14, Amsterdam: North–Holland.

[10] Mittnik, S. & S.T. Rachev (1993), Modeling asset returns with alternative stable distribution. *Econometric Review* 12, 261-330.

[11] Myers R.H., I. Khuri & W.-H. Carter, Jr. (1989), Response surface methodology: 1966–1988. *Technometrics* 31, 137–157.

[12] Neyman, J. & E.L. Scott (1971), Outlier proneness of phenomena and of related distributions. In J.S. Rustagi ed., *Optimizing Methods in Statistics*, New York: Academic Press.

[13] Rachev, S.T. & S. Mittnik (2000), *Stable Paretian Modeling in Finance.* New York: Wiley.

[14] Rachev, S.T., S. Mittnik & J.-R. Kim (1996), Outlier detection in heavy–tailed samples, manuscript Department of Statistics, University of California, Santa Barbara.

[15] Resnick, S.I. (1987), *Extreme Values, Regular Variation, and Point Processes,* New York: Springer.

[16] Samorodnitsky G. & M.S. Taqqu (1994), *Stable Non-Gaussian Random Processes*, New York: Chapman & Hall.

[17] Weron, R. (1996), On the Chambers–Mallows–Stuck method for simulating skewed stable random variables. *Statistics and Probability Letters* 28, 165–171.

[18] Zolotarev, V.M. (1986), *One-dimensional Stable Distributions*, Translations of Mathematical Monographs, American Mathematical Society, Vol. 65, Providence.