A Contemporary Review and Bibliography of Infinitely Divisible Distributions *

by

Arup Bose and Anirban DasGupta
Indian Statistical Institute, Purdue University

Herman Rubin
Purdue University

Technical Report #01-13

Department of Statistics
Purdue University

October 2001

* Invited Article for the Golden Jubilee Volume, Indian Institute of Technology

A Contemporary Review and Bibliography of Infinitely Divisible Distributions

Arup Bose
Indian Statistical Institute, Calcutta

Anirban DasGupta
Purdue University

Herman Rubin
Purdue University

This article gives a modern review of univariate and multivariate stable and infinitely divisible distributions. Various characterizations and properties of stable and infinitely divisible distributions, including tail, moment and independence properties, and methods of simulation from an infinitely divisible distribution are discussed. Also discussed is the currently popular problem of estimating the index of a stable law and more generally, the heaviness of the tail of a distribution in the domain of attraction of a given stable law. A special feature of this article is its large collection of illustrative examples.

AMS *Subject Classifications*. Primary 60E07, Secondary 60-02, 62F10

Key words. Infinitely divisible, stable, characteristic function, Levy measure, spectral measure, domain of attraction, compound Poisson, completely monotone, tail, moment, Hill estimate.

1. Introduction

Infinitely divisible distributions were introduced by de Finetti in 1929 and the most fundamental results were developed by Kolmogorov, Lévy and Khintchine in the thirties. The area has since continued to flourish and a huge body of deep and elegant results now exists in the literature. There have been many significant developments in the area in the last 20 to 25 years, and a contemporary review seems to be needed. This article provides a review on both univariate and multivariate infinitely divisible distributions, with a significant review of the recent development in inference, simulation and applications.

The first question is what are infinitely divisible (id) distributions? The following definition most fits the name, although other equivalent characterizations are available and are to be described later.

Definition 1. A real valued random variable X with cumulative distribution function (cdf) $F(\cdot)$ and characteristic function (cf) ϕ is said to be infinitely divisible (id), synonymously F is an id law or ϕ is id, if for each $n > 1, X$ can be divided into n independent and identically distributed (iid) components, i.e., there exist iid random variables $X_1, \ldots, X_n$ with cdf say F_n such that X has the same distribution as $X_1 + \ldots + X_n$.

Remark 1. Since such a "division" of X into "small" independent components is possible for each n, the name infinitely divisible seems appropriate.

Example 1. Let X have the standard normal distribution. For a given n, take $X_1, \ldots, X_n$ as iid $N(0, 1/n)$. Then X has the same distribution as $X_1 + \ldots + X_n$. Thus the standard normal distribution is id. Indeed, this argument shows all univariate normal distributions to be id.

Example 2. Let X have a Poisson distribution with mean 1. For a given n, take $X_1, \ldots, X_n$ as iid Poisson random variables with mean $1/n$. Then X has the same distribution as $X_1 + \ldots + X_n$. Thus the Poisson distribution with mean 1 is id. Indeed, this argument shows all Poisson distributions to be id.

Example 3. Let X have the continuous uniform $[0, 1]$ distribution. Then X is *not* id. For if it is, then for any n, we will be able to find iid random variables $X_1, \ldots, X_n$ with some distribution F_n such that X has the same distribution as $X_1 + \ldots + X_n$. Since X takes values in $[0, 1]$, it will force the supremum of the support of F_n to be at most $1/n$. In turn, this will force the variance of X_n to be at most $1/n^2$ and so the variance of X to be at most $1/n$, which would patently contradict the fact that the variance of X is $1/12$. We see that X cannot be id.

Remark 2. Indeed, any random variable X with a bounded support cannot be infinitely divisible, unless of course X is a constant. The method outlined in Example 3 works for all such X. As a consequence, binomial, hypergeometric, and beta distributions are not id.

This raises the natural question:

Question 1. Which real valued random variables with unbounded support are id?

This question can be completely answered via several equivalent characterizations available for id laws. Some of these are given below. Interestingly most common univariate random variables with unbounded support *are* infinitely divisible. But there are a few common univariate random variables with unbounded support that are *not* infinitely divisible. Here are two lists which cover univariate distributions in common use:

List 1. Those that are infinitely divisible: This includes the discrete distributions such as Poisson, geometric and negative binomial and the continuous distributions normal, lognormal, noncentral chi-square, t, exponential, Gamma, double exponential, Pareto, Cauchy, half Cauchy and squared Cauchy.

List 2: Those that are not infinitely divisible: This includes finite mixtures of normals, discrete normal, absolute normal, inverse normal and inverse t.

Remark 3. In a good number of these cases, the proof that a certain distribution is or is not id nontrivial and an utterly specialized task. Instances of this are the proofs that half Cauchy, lognormal, and t distributions are id. Similarly, the proofs that inverse normal and inverse t distributions are not id require tricks that are not well known. It is a peculiarity of the subject, a bit like admissibility in decision theory, that hard special techniques may be needed for particular special problems.

2. Characterizations

Now let us return to Question 1. A number of equivalent characterizations will be given. We shall also discuss results known for subclasses such as the class of all nonnegative random variables and all nonnegative random variables which have density.

2.1 General characterizations.

First, let us see another familiar but motivating example.

Example 4. Fix $n > 1$, and take $X_{n1}, \ldots, X_{nn}$ to be iid Bernoulli (p_n) random variables. Then $S_n = X_{n1} + X_{n2} + \ldots + X_{nn}$ has a binomial(n, p_n) distribution, and if $p_n \to 0$ as $n \to \infty$ in such a way that $np_n \to \lambda$, for some $0 < \lambda < \infty$, then S_n converges in distribution to a Poisson random variable with mean λ which is id.

Note that the distribution of X_{in} does depend on n, and that the limit distribution is Poisson, which is id. Hence we may ask the following question:

Question 2. Fix $n \geq 1$. Take $X_{n1}, X_{n2}, \ldots, X_{nn}$ to be iid with some common distribution say H_n. Take as in Example 4, $S_n = X_{n1} + X_{n2} + \ldots + X_n$. If S_n has a limit distribution, say F, can we assert anything interesting about the nature of F?

The answer provides our first characterization:

Characterization # 1: Such an F is id and conversely, every id F arises in this fashion.

Remark 4. In Question 2, the random variables $X_{n1}, \ldots, X_{nn}$ at the nth stage have a distribution H_n that depends on n; in other words, the random variables X_{in} formed a *triangular array.* Suppose instead that we have one sequence of summands which are iid but we allow appropriate centering and normalization. This raises the following question:

Question 3. Suppose $X_1, X_2, \ldots$ is an iid sequence with some common distribution H. Take $S_n = X_1 + X_2 + \ldots + X_n$; suppose, for some sequences of numbers a_n and b_n, $(S_n - a_n)/b_n$ has a limit distribution, say F. Can we assert anything interesting about the nature of F?

Question 3 is a special case of Question 2 by taking $X_{in} = X_i/b_n - a_n/(nb_n)$. So certainly our limit law F in Question 3 is id. The collection of all such $F's$ is thus a subclass of the class of all id laws. This subclass is the class of all *stable laws* and can also be defined as follows:

Definition 2. A cdf F on the real line is said to be stable if for every $n \geq 1$, there exists constants b_n and a_n such that $S_n = X_1 + X_2 + \ldots + X_n$ and $b_n X_1 + a_n$ have the same law. Here $X_1, X_2, \ldots, X_n$ are iid with common distribution F.

Remark 5. Thus F is stable if the sum of n iid observations from F has the same *type* of distribution as one observation X_1. By same type we simply mean that a location–scale transformation of X_1 gives the distribution of the sum S_n. It turns out that b_n has to be asymptotically equivalent to $n^{1/\alpha}$ for some $0 < \alpha \leq 2$. The constant α is said to be the *index* of the stable distribution F.

Example 5. It is known that all stable laws have densities that are infinitely differentiable and all derivatives are bounded. However, there are only three types of stable laws for which the density functions are known in simple closed forms. These are the (i) the normal distribution which is stable with index 2, (ii) the Cauchy distribution which is stable with index 1 and (iii) the *Levy distribution* which is stable with index 1/2.

Let us return to the issue of characterizing id laws at large. There is a very elegant characterization of id laws as *compound Poisson* distributions.

Characterization # 2. Take an infinite sequence of iid random variables $X_1, X_2, \ldots$, with a distribution say H, and take a Poisson random variable N independent of the X_i's. Define a new random variable X as

$$X = X_1 + \ldots + X_N;$$

then X is infinitely divisible. Conversely, every id law arises in this way.

A nice use of this characterization is the following:

Example 6. Take X to have a noncentral chi–square distribution with say one degree of freedom and some noncentrality parameter. It is well known that X may be written as

$Y_1 + Y_2 + \ldots + Y_{2N+1}$, where the Y_i are iid central chi-squares with one degree of freedom and N is an independent Poisson random variable. Write $X_1 = Y_2 + Y_3, X_2 = Y_4 + Y_5$, etc. Then $X = Y_1 + (X_1 + X_2 + \ldots + X_N)$, where Y_1 is id. The quantity in parentheses is also id by characterization # 2, and hence so their sum. That is, X is id.

We next turn to the most common means of characterization of id laws, namely by their characteristic functions. These characterizations are known in a number of different forms. Some are easier to describe, while others are easier to apply. We give two of these forms, one for the finite variance case, and another for the general case.

Characterization # 3.

Form A Let F be an id law with mean b and finite variance and let $\phi(t)$ denote its characteristic function. Then $w(t) = \log\phi(t)$ admits the representation

$$w(t) = ibt + \int_{-\infty}^{\infty}(e^{itx} - 1 - itx)\frac{d\mu(x)}{x^2} \tag{1}$$

where μ is a finite measure on the real line. Furthermore, $\mu(\mathbb{R}) = Var(X)$.

Example 7. Suppose F is the normal distribution with mean 0 and variance σ^2. Then the measure μ is degenerate at 0 with point mass σ^2 there and b is 0.

Example 8. Suppose Y has a Poisson distribution with mean λ and take F to be the distribution of a location – scale transformed version $X = c(Y - \lambda)$. Then the measure μ is degenerate with mass c^2 at c and again, b is 0.

Remark 6. Form A can be roughly interpreted as follows. The measure μ corresponding to a general infinitely divisible law F is degenerate for the cases when F is normal or Poisson. A general μ can be approximated by linear combinations of such degenerate measures. In other words, a general infinitely divisible law is a limit of finite convolutions of normal and Poisson type random variables. This is in fact true without the finite variance assumption made in describing Form A above.

Form B Let F be an id law and let $\phi(t)$ denote its characteristic function. Then $w(t) = \log\phi(t)$ admits the representation

$$w(t) = ibt - t^2\sigma^2/2 + \int_{-\infty}^{\infty}(e^{itx} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\lambda(x) \tag{2}$$

where b is a real number, and λ is a finite measure on the real line giving mass 0 to the value 0, i.e., $\lambda\{0\} = 0$. The integrand is defined to be $-t^2/2$ at the origin, by continuity.

Remark 7. This is the original canonical representation for the characteristic function of an infinitely divisible law given by Paul Lévy. However, for certain applications and special cases, Form A is more useful.

We have seen that a Poisson random variable X is id. A *Poisson type* random variable is one of the form $aX + b$. Id laws can also be characterized by writing them as limits of convolutions of such *Poisson type* random variables.

Characterization # 4. F is infinitely divisible if and only if it is the limit in distribution of $S_n = X_1 + X_2 + \ldots + X_n$, where X_i are independent *Poisson type* random variables.

Remark 8. This representation of infinitely divisible distributions has been used to simulate observations from an infinitely divisible law. See section 5 later.

2.2 Characterization of nonnegative discrete id laws.

There is an elegant characterization of all distributions supported on the nonnegative integers that are id. The characterization says the following.

Let X take values $0, 1, 2, \ldots$, with $P(X = k) = p_k$. Then X is id if and only if

$$\eta_i = \frac{ip_i}{p_0} - \sum_{j=1}^{i-1} \eta_{i-j} \frac{p_j}{p_0} \geq 0 \quad \forall i \geq 1.$$

Remark 9. Due to the recursive nature of this characterization, it is usually difficult to use it for verifying that some distribution is id. But it can be used to verify that a certain distribution is not id. Let us see an example.

Example 9. Consider the discrete standard normal distribution with mass function $p_k = 2[\theta(2\pi)^{-1}) + 1]^{-1} \exp(-k^2/2)$ where $\theta(\cdot)$ is the Jacobi theta function. Then $\eta_1 = .6065 > 0$, but $\eta_2 = -.0972 < 0$, and it follows that the discrete normals are not infinitely divisible.

However, simple and verifiable sufficient conditions that imply the above characterization are available. One such sufficient condition is the following:

Sufficient Condition. Let X take values $0, 1, 2, \ldots$, with $P(X = k) = p_k$. Then X is id if $\log p_k$ is convex in k.

This is often verifiable. Another very nice consequence of the characterization result is a *necessary* condition for a discrete distribution to be id. It says the following:

Necessary Condition. Let X take values $0, 1, 2, \ldots$, with $P(X = k) = p_k$, with $p_1 > 0$. Suppose X is id; then for all $k > 1$, $p_k > 0$.

Remark 10. This says that the support of a discrete id law cannot have any gaps if $P(X = 1)$ is strictly positive. We will see below that a similar result holds for positive id random variables with a density as well.

2.3 Nonnegative id laws with densities.

One of the most important and useful results on id laws is the Goldie–Steutel law for positive random variables, which we will describe shortly. In general, the proof that a certain positive random variable having a density is id may be accomplished in one of the following ways:

a. verify that the Goldie–Steutel law applies;

b. try to verify a known characterization result parallel to the discrete case;

c. use a known sufficient condition;

d. use a special technique for that particular problem.

The problems that are solved by using technique d) generally turn out to be the difficult ones; the lognormal and the half Cauchy are two instances.

The Goldie–Steutel Law. Let a positive random variable X have a density $f(x)$ which is *completely monotone.* Then X is id.

Remark 11. *Complete monotonicity* means that the function is infinitely differentiable, decreasing, and derivatives of successive orders have opposite signs. It is well known that such functions may be written as exponential mixtures. So the Goldie–Steutel Law may also be stated as saying that a positive random variable X with density f is id if it can be written as $X = YZ$, where Z is exponential with mean 1, and Y is nonnegative and independent of Z. The nice thing about the Goldie–Steutel Law is that many positive random variables are known to be of this variety, and a fortiori, they are id.

Example 10. Let X have the Pareto density $f(x) = \frac{\alpha}{\mu}(\frac{\mu}{x+\mu})^{\alpha+1}$. Then, easily, f is completely monotone, and so X is infinitely divisible. It can be verified that in the representation $X = YZ$ as above, Y has a Gamma density.

There is an extension of the Goldie–Steutel law that is sometimes useful. The extension says that certain *improper* exponential mixtures are also infinitely divisible. The statement is as follows:

Extension of the Goldie–Steutel Law. Let X have a density $f(x)$ of the form

$$f(x) = \int_0^\infty \exp(-xt)g(t)dt,$$

where $g(\cdot)$ changes sign once. Then X is id.

Now let us see the continuous analogs of some of the results we saw in the discrete case. First, a characterization of id laws in terms of the density function:

Characterization of nonnegative id laws with densities. A nonnegative random variable $X > 0$ with density function $f(x)$ is id *if and only if* there is a nondecreasing function $\tau(u)$ on $[0, \infty]$ with $\int_1^\infty u^{-1}d\tau(u) < \infty$, such that $f(x)$ may be written as $f(x) = x^{-1}\int_0^x f(x-u)d\tau(u)$.

Verifying whether a given f may be written in the above form corresponds to solving an integral equation with difficult constraints. So as in the discrete case, the above characterization is not necessarily easy to apply. But fortunately, there are certain verifiable sufficient conditions and necessary conditions for applications. We give one pair below.

Sufficient condition. Let X be a positive random variable with a strictly positive decreasing and twice continuously differentiable density $f(x)$. Then X is id if

$$\frac{f'(y)}{f(y)} \leq \frac{1}{x} + \frac{f''(x)}{f'(x)}, \quad \forall\, 0 < y \leq x.$$

This is very explicit and one can attempt to analytically verify it. Another extremely elegant necessary condition is the following.

Necessary condition. Let X be a positive id random variable with a density $f(x)$. If $f > 0$ in some neighborhood of 0, then it cannot have any zeroes.

Remark 12. Note that this parallels the result for the discrete case that the support of X cannot have any gaps.

3. Properties of id laws

Id laws have very interesting properties in terms of their characteristic function, moments and tails. Moreover, subclasses of id laws such as those that are unimodal, totally positive etc turn out to be quite interesting. We shall discuss some of these below. The important subclass of *stable laws* is treated separately in subsection 3.3.

3.1 Properties of the characteristic function.

Characteristic functions of id laws satisfy some clean properties. Such properties are useful to exclude particular distributions from being id and to establish further properties of id laws as well. They generally do not provide much probabilistic insight, but are quite valuable as analytical tools in studying id laws. A collection of properties is listed below.

1. Let $\phi(t)$ be the characteristic function (cf) of an id distribution. Then ϕ has no real zeroes. The converse is false.

2. Let $\phi(t)$ be the characteristic function (cf) of an id distribution. Then for all $\lambda > 0$, $\phi^\lambda(t)$ is also a cf. Here $\phi^\lambda(t)$ is to be defined as $\exp(\lambda Log[\phi(t)])$, where $Log[.]$ denotes the *distinguished logarithm*.

3. Let $\phi_1(t), \phi_2(t)$ be two id cfs; then $\phi_1(t)\phi_2(t)$ is also an id cf.

4. Let $\phi(t)$ be the characteristic function (cf) of an id distribution. Then $\overline{\phi(t)}$, the complex conjugate of ϕ, and $|\phi|^2(t)$ are also id cfs.

5. Let $\phi_n(t)$ be a sequence of id cfs, converging pointwise to another cf $\phi(t)$. Then $\phi(t)$ is also an id cf.

6. Let $\phi(t)$ be the characteristic function (cf) of an id distribution. Then there exist real constants a, b, such that $|\log \phi(t)| \leq a + bt^2$ for all t.

Remark 13. Proofs of properties 1 and 2 can be found in any advanced text on probability. Property 3 just says that the convolution of id laws is id. Property 4 says that the negative of an id random variable X is id, and therefore if X_1 and X_2 are iid and id, then $(X_1 - X_2)$ must also be id. Property 5 is essentially a restatement of Characterization # 1. Let us see a quick example that shows that the converse of Property 1 is false.

Example 11. Consider the function $\phi(t) = (\cos t + 2)/3$. This is the cf of a symmetric distribution supported on $\{-1, 0, 1\}$ and obviously has no real roots. And also, evidently, this random variable cannot be id.

3.2 Moments and tails of id laws.

An id random variable may have all moments, some moments, or even no moments. For instance, the normal has all moments, the Cauchy has no moments, and intermediate t distributions have some moments. But one can say some definite things about the tails of id laws. For example, roughly speaking, no id law can have tails thinner than that of a normal. We state these and connections to the canonical measures of id laws below.

Let X be an infinitely divisible random variable with cdf $F(x)$ and corresponding canonical measure λ as in Form B of its characteristic function. Then,

1. $-\log(1 - F(x) + F(-x)) = O(x \log x)$ as $x \to \infty$ unless F is a normal cdf;

2. There cannot exist any reals $a > 0$, $b > 1$ such that $1 - F(x) + F(-x) = O(\exp(-ax^{1+b}))$ as $x \to \infty$ unless F is degenerate;

3. There cannot exist any reals $a > 0, 0 < b \leq 1$ such that $1 - F(x) + F(-x) = O(\exp(-ax^{1+b}))$ as $x \to \infty$ unless F is normal;

4. If $\lim_{x \to -\infty} F(x)/\Phi(x) = 1$, then F must be Φ itself;

5. For a given $p > 0, 1 - F(x) = O(x^{-p})$ as $x \to \infty$ if and only if $\int_x^\infty d\lambda(u) = O(x^{-p})$ as $x \to \infty$; and $F(-x) = O(x^{-p})$ as $x \to \infty$ if and only if $\int_{-\infty}^{-x} d\lambda(u) = O(x^{-p})$ as $x \to \infty$.

6. For a given $p > 0, E(|X|^p) < \infty$ if and only if $\int_{-\infty}^\infty |u|^p d\lambda(u) < \infty$.

Remark 14. The connections of F to the canonical measure λ via their respective tails as in 5) are nice; so is the equivalence between existence of absolute moments. It is also interesting that the assertion of 4) is false if the cdf in the denominator is an id cdf $G(x)$ other than $\Phi(x)$.

Example 12. Suppose X has the density function $f(x) = 1/(2\Gamma(5/4)) \exp[-x^4]$. Then, from 1) or 2), it follows that X cannot be infinitely divisible. The tail of $f(x)$ is too thin for X to be infinitely divisible.

Or, suppose X has a mixture normal distribution $pN(0,\sigma_1^2) + (1-p)N(0,\sigma_2^2)$, for unequal σ_1^2, σ_2^2. Then, from 1), it follows that X cannot be id.

3.3 The stable laws

The subclass of stable laws occupies a special position in the class of id laws. Their probabilistic properties have been studied extensively. They have also found numerous applications in statistics. In this subsection we discuss some of the probabilistic properties of stable laws. In section 6 we shall look at the statistical importance of this class.

Characteristic function of stable laws: Starting from Form B of the characteristic function of id laws, it is possible to derive the following characterization:

$\phi(t)$ is the cf of a stable law F:

with index $\alpha \neq 1$ if and only if it has the representation

$$\log \phi(t) = ibt - \sigma^\alpha |t|^\alpha (1 - i\beta sign(t) \tan \frac{\pi\alpha}{2}) \tag{3}$$

with index $\alpha = 1$, if and only if it has the representation

$$\log \phi(t) = ibt - \sigma |t| (1 + i\beta sign(t) \frac{2}{\pi} \log |t|).$$

The *scale parameter* $\sigma > 0$, the *location parameter* b and the *skewness* parameter β of F above are *unique.* The possible value of β ranges in the closed interval $[-1,\ 1]$. The possible values of b are the entire real line. It follows trivially from the characteristic function that F is *symmetric* (about b), if and only if $\beta = 0$. If $\alpha = 2$ then F is normal. In this case, the value of β is irrelevant. If $\alpha = 1$ and $\beta = 0$, then F is a Cauchy law with scale σ and location b.

Moments and tails of a stable law. The stable laws also have some very nice moment and tail and properties. But first an easy fact:

The first moment of any random variable, if it exists, is equal to the first derivative of the cf at zero. Thus by using the above characterization, it is easy to see that if X is stable with $\alpha > 1$, then $E(X) = b$.

What can be said about other moments? Of course, if $\alpha = 2$, then all moments exist.

Moment behavior: If X is stable with $0 < \alpha < 2$, then for any $p > 0$,

$$E|X|^p < \infty \text{ if and only if } 0 < p < \alpha$$

This property of the moments *suggests* that the tails of a stable law behave as $x^{-\alpha}$. This is essentially correct:

Tail behavior: If X is stable with index $0 < \alpha < 2$, then there exists a non zero constant $C_\alpha \neq 0$, such that,

$$\lim_{x \to \infty} x^\alpha P\{X > x\} = C_\alpha (1+\beta)\sigma^\alpha/2$$

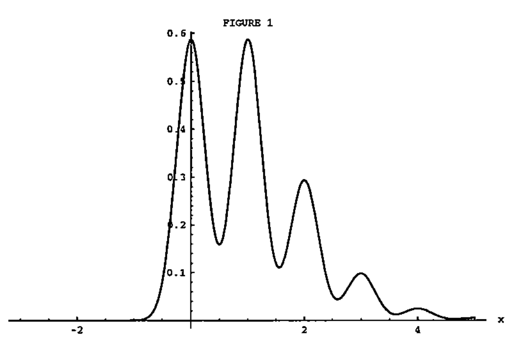
$$\lim_{x \to \infty} x^\alpha P\{X < x\} = C_\alpha (1-\beta)\sigma^\alpha/2$$

Thus a stable law of index $0 < \alpha < 2$ has *at least* one of the tails of *exact* asymptotic order $x^{-\alpha}$. If $\beta \neq 1, \ -1$, then *both* tails are of this order.

3.4 Unimodality, total positivity and the class L.

The standard examples of id laws are all unimodal. Even the discrete ones are discrete unimodal. But it is not difficult to construct simple continuous id random variables which do not have unimodal densities. Let us see an example.

Example 13. Suppose X_1 has the $N(0, \sigma^2)$ distribution, and X_2, independent of X_1, has a Poisson (λ) distribution. Consider the convolution $X = X_1 + X_2$. Evidently, X is infinitely divisible. However, for given λ, the density of X will not be unimodal for sufficiently small σ. Figure 1 gives the density of X when $\lambda = 1$ and $\sigma = 1/4$. An X of this form, in general, has a density with finitely many distinct local maxima.



In view of this, it is interesting to ask what can be said about unimodality of id laws. It turns out that a large class of id laws having a certain property known as *self decomposability* are indeed unimodal. We first give the definition.

Definition 3. Let X be a random variable with characteristic function $\phi(t)$. X is said to be *self decomposable* if for every $0 < c < 1, \phi(t)$ can be factorized as $\phi(t) = \phi(ct)\psi(t)$, where ψ is another characteristic function. In other words, X can be written as a convolution $X \overset{\mathcal{D}}{=} cX + Y$ for every $0 < c < 1$. The class of all such laws is called *the class L*.

Remark 15. Clearly, normal and Cauchy distributions are self decomposable. We state a more general result below in 2).

1. Every id random variable that is in the class L is unimodal.

2. All stable distributions belong to the class L.

3. All stable distributions, which are a fortiori id, are unimodal.

Remark 16. Thus, we have a nice subclass of id laws,namely all stable laws, that are unimodal. The proof that every density in the class L is unimodal is nontrivial. Note that the result stated above indicates how to construct other large subclasses of id laws that would also be unimodal. For instance, take a convolution of a normal random variable with a stable random variable. This will be id, and will also be unimodal because a normal random variable is *strongly unimodal* and a stable one, as we just stated, is unimodal.

For one sided, i.e., either positive or negative, stable laws, it is sometimes possible to assert a very strong kind of unimodality. It is the following :

Let X be a positive stable random variable with index $\alpha = 1/k$ for some natural number k and $|\beta| = 1$ in the canonical representation of its characteristic function. Then the density of X is *totally positive.*

3.5. Approximation of sums in total variation by id laws.

Take iid random variables $X_1, X_2, \ldots$ having some common distribution H. Under well known conditions, $S_n = X_1 + X_2 + \ldots + X_n$, when centered and normalized, will converge to a normal distribution. However, the convergence is not necessarily in total variation. A simple example is that of iid Bernoulli random variables X_n. In this case, for all n, the total variation distance remains equal to 1. However, if H is continuous with a unimodal density, then the convergence will also be in total variation (in fact a more general result is true outside of this narrow central limit structure).

So it is interesting that if the approximating class is enlarged to the class of id laws, then it is possible to say definite things about convergence in total variation. A few results are as follows:

Let $X_1, X_2, \ldots$ be an iid sequence with some common distribution H. Let $\mathcal{I}$ denote the class of all id laws and let $H_n(x) = P(S_n \leq x)$. Then

1. $\lim_{n\to\infty} \sup_{H} \inf_{F \epsilon \mathcal{I}} d_{TV}(H_n, F) = 0;$

2. There exist finite constants c_1, c_2, such that

$$c_1(n \log n)^{-1} < \sup_H \inf_{F \epsilon \mathcal{I}} d_{TV}(H_n, F) < c_2 n^{-1/3} (\log n)^2;$$

3. If X_j are iid Bernoulli (p), then

$$\sup_{0<p<1} \inf_{F \epsilon \mathcal{I}} d_{TV}(H_n, F) = O(n^{-2/3}).$$

4. Multivariate id laws

Let X be a k–dimensional random vector. Then the definition of infinite divisibility of X is the same as in one dimension. Many of the results are similar too, for instance inclusion of the multivariate stable laws, canonical representations of the characteristic functions, etc. The basic theorems about id random variables were generalized to id vectors as early as 1954. Early references in this area are Rvaĉeva (1954), Takano (1954) and Dwass and Teicher (1957).

Of course if a random vector is id, then all the lower dimensional components are also so. However, interesting things happen when we consider other aspects. In the subsections below, we discuss some aspects of multivariate id laws, such as, independence, Gaussianity, existence of moments etc.

4.1. Some interesting examples.

We will report here a collection of results and examples which explore the connection between the concepts of infinite divisibility of the full vector and lower dimensional transformations.

Example 14. It is possible that a random vector X is not id, but all linear combinations of the coordinates of X are id. Here is a somewhat natural example. Let Z be a standard bivariate normal vector. Define a new bivariate random vector $X = (c'Z, Z'AZ)$, where c is a 2–tuple and A is a 2×2 symmetric matrix. If c is not in the null space of A, the vector X is not id. However, every linear combination of the two coordinates of X is infinitely divisible.

Example 15. It is possible that every lower dimensional projection of a multivariate random vector X is id, while X itself is not id. Towards this end, take Z_1, Z_2 to be iid $N(0, 1)$, and define a new trivariate vector X as $X = (Z_1^2, Z_1 Z_2, Z_2^2)$. Then, it is easily verified that each two dimensional projection (and hence, each one dimensional projection as well) is id, but X itself is not id.

Example 16. For iid univariate normal observations, the sample variance is a scaled chi–square and hence id. Curiously, the corresponding result for the multivariate case is not true. For simplicity, let us assume that we have a nonsingular normal distribution. Thus, suppose we have iid observations from a k–dimensional normal distribution, and suppose S is the usual Wishart matrix of sample variances and covariances. Then, S is not id.

Example 17. Let X be an id random vector. Then it is possible that although X is not multivariate normal, certain linear combinations $c'X$ of X are univariate normal. Indeed, $c'X$ is univariate normal if and only if the Lévy measure corresponding to the distribution of X is supported on the manifold $\{x : c'x = 0\}$. Of course, such examples of normal projections of nonnormal vectors are well known; but now the full vector itself is id.

Example 18. This example shows that one can have a bivariate random vector that is not id, but the product of its coordinates is id. It thus gives a nonlinear function of the coordinates that is id, while linear functions were considered in some of the preceding examples. Towards this, take a $N(0,1)$ random variable Z and write Z as $Z = X_1X_2$, where X_1, X_2 are iid. This is possible. Take the random vector X to be (X_1, X_2). Now, clearly, X_1X_2 is id by construction, but X is not infinitely divisible. To show that X is not infinitely divisible, it is enough to show that X_1 is not id. The reason for this is that X_1 has too thin a tail. Indeed, if $|X_1|, |X_2|$ are each larger than some x, then $|Z|$ is larger than x^2; hence,

$$P(|X_1| > x) \leq P(|Z| > x^2)^{1/2} = O(x^{-1}\exp(-x^4/2)),$$

and from the first fact 1) listed under moments and tails of id laws (section 3.2), we see that X cannot be id.

4.2 When are components of an id vector independent?

Recall the universally known fact about Gaussian random vectors: if $(X_1, \ldots, X_n)$ is Gaussian then the components are mutually independent if and only if they are pairwise independent which in turn happens if and only if $\mathrm{Cov}(X_i, X_j) = 0 \ \forall \ i \neq j$.

Now assume that $X = (X_1, \ldots, X_n)$ is id. A natural question is when are the components independent? Are there any necessary and sufficient conditions available as in the normal case?

It turns out that if the id vector has finite fourth moment, then pairwise independence is still equivalent to total independence. Thus if we restrict to id vectors with finite fourth moment, then the problem reduces to that of finding conditions for pairwise independence.

Since an id vector can, in general, have Poisson components, it is clear that the covariance condition which is necessary and sufficient for pairwise/total independence when X is normal does not remain so when X is merely id. But interestingly, the addition of one extra condition leads to a satisfactory solution. Assume that X is id and has a finite fourth moment. Since total independence is equivalent to pairwise independence, it is enough to concentrate on the case where X is a 2-vector, $X = (X_1, X_2)$. To simplify expressions, assume that $E(X_i) = 0$ for $i = 1,\ 2$.

Let

$$\beta = (2,2) \text{ cumulant of } (X_1, X_2).$$

This cumulant is given by:

$$\beta = \mathrm{Cov}(X_1^2,\ X_2^2) - 2(\mathrm{Cov}(X_1, X_2))^2.$$

From the results of Pierre (1971) (see also Sclove (1981)), it is known that $\beta \geq 0$.

In general, the two components X_1 and X_2 are independent if and only if $\text{Cov}((X_1,\ X_2) = 0$ *and* $\text{Cov}((X_1^2,\ X_2^2)) = 0$.

In several special cases, β carries the information on independence of the components. For example, if (X_1, X_2) has *no* Gaussian component, then X_1 and X_2 are independent if and only if $\beta = 0$. In particular, if X is discrete then X_1 and X_2 are independent if and only if $\beta = 0$. It also follows that in general, the Poisson components are independent if and only if $\beta = 0$.

4.3. When are id vectors Gaussian?

Suppose $X = (X_1, \ldots, X_n)$ is id. As we have discussed in Example 17, it is possible that *certain* linear combinations are normal but X is not normal. What happens if sufficiently many linear combinations are normal? Indeed, if each X_k is Gaussian, then X is Gaussian. One can say more. If there is at least one component k such that the 4th cumulant of X_k is zero then also X is Gaussian.

Recall that the regression functions of Gaussian vectors are linear. Further, all conditional distributions are homoscedastic. That is, the dispersion matrix of any subvector given any other is free of the conditioning subvector. For characterization of normal vectors using such ideas, see Kagan et. al. (1973).

However, for id X, homoscedasticity and the linearity for vectors up to a pair guarantees Gaussianity of X, We can state the following precise result.

Suppose X is square integrable, linearly independent (to avoid trivialities) and with pairwise nonzero correlations. Suppose for some $i,\ j, k$,

$$
\begin{aligned}
E(X_i|X_j) &= a_{ij}X_j + \mu_{ij} \\
Var(X_i|X_j) &= b_{ij} \\
E(X_i|X_j, X_k) &= a_{j,(i,j,k)}X_j + a_{k,(i,j,k)}X_k + \alpha_{i,(j,k)} \\
Var(X_i|X_j, X_k) &= b_{i,jk}
\end{aligned}
$$

Then X is Gaussian. For more information on such characterizations, see also Wesolowski (1993) and Arnold and Wesolowski (1997).

4.4 Multivariate stable distributions.

Multivariate stable laws forms a subclass of multivariate id laws. While they have not found much applications in statistical modelling yet, it is believed that this situation will change in the near future. In particular, they are anticipated to be of much use in economic data modelling. There are different ways of extending the univariate notion of stability, giving rise to different classes of multivariate stable laws. We will take the following as our definition:

Definition 4. A random vector $X = (X_1, \ldots, X_k)$ with distribution F is said to be stable, equivalently F is said to be stable if for independent copies $X^{(1)}$ and $X^{(2)}$ of X, and for any positive numbers a and b, there exists a positive number c and a vector D such that $aX^{(1)} + bX^{(2)} \stackrel{\mathcal{D}}{=} cX + D$. If $D = 0$, then X is said to be *strictly stable.*

As in the univariate case, if X is stable, there is an α, $0 < \alpha \leq 2$, called the index of X or F, such that for any $n \geq 2$, there is a vector D_n such that $X^{(1)} + X^{(2)} + \ldots + X^{(n)} \stackrel{\mathcal{D}}{=} n^{1/\alpha}X + D_n$ where $X^{(1)}, \ldots X^{(n)}$ are iid copies of X. Moreover, this can be taken as the definition of multivariate stability, equivalent to the one given above.

Example 19. Of course, as in the univariate case, if $\alpha = 2$, then X is multivariate normal.

Example 20. It is not hard to verify the following: If X is α stable (respc. strictly stable) then *all* linear combinations are α stable (respc. strictly stable).

What about the converse? It turns out that the converse is *partially* true and we have the following facts:

1) If *all* linear combinations of the coordinates of X are stable with $\alpha \geq 1$, then X is stable.

2) If *all* linear combinations of the coordinates of X are *strictly stable*, then X is strictly stable.

3) If *all* linear combinations of the coordinates of X are *symmetric stable*, then X is symmetric stable. (Here symmetry is defined as $X \stackrel{\mathcal{D}}{=} -X$).

Example 21. The conclusion in 1) is *false* in general if $0 < \alpha < 1$. To see this take $\Psi(t_1, t_2) = \exp\{-r^\alpha + i\rho r\cos(3\phi)\}$ where $t_1 = r\cos(\phi)$, $t_2 = r\sin(\phi)$. Then for sufficiently small $\rho > 0$, Ψ is a characteristic function of a vector X which is *not* stable. However, it is rather easy to check via the characteristic function that *any* linear combination of the two coordinates of X is stable.

Remark 17. Actually, in the above example, X is not even id. In general, if we assume that X is id and all linear combinations are stable then X is also stable.

The spectral measure of a stable law. If X is stable with $0 < \alpha < 2$, then its characteristic function has the following representation. This representation can be arrived at starting from the representation of id laws.

Let S denote the unit sphere in k dimensions and Γ a finite measure on S. Then, with $<,>$ denoting inner product, the cf of a stable law has the representation

$$\Psi(t) = \exp\{i < t,\ \mu > - \int_S | < t, s > |^\alpha \times [1 - isign(< t,\ s > \tan\frac{\pi\alpha}{2})]\Gamma(ds)\}, \qquad (4)$$

if $\alpha \neq 1$.

If $\alpha = 1$, then

$$\Psi(t) = \exp\{i < t,\ \mu > - \int_S |< t, s >| \times [1 + i\frac{2}{\pi} sign(< t,\ s > \log|< t,\ s >|]\Gamma(ds)\} \quad (5)$$

The pair (Γ, μ) is unique. The above representation is called the *spectral representation.* Γ is called the *spectral measure.*

Example 22. The characteristic function of the *multivariate Cauchy* random variable X is given by

$$\Psi(t) = \exp\{-(t'\Sigma t)^{1/2} + i < t,\ \mu >\}$$

If Σ is the identity matrix and $\mu = 0$, then X is spherically symmetric stable with Γ being the *uniform measure.* Its density is given by

$$f(x) = (2\pi)^{-1}(1 + x_1^2 + x_2^2)^{-3/2} \quad -\infty < x_1,\ x_2 < \infty$$

Example 23. From the above representation, we can derive a criterion for the independence of the components of a stable vector. If X is stable, then its components are independent if and only if the spectral measure Γ is discrete and is concentrated on the intersection of the axes with the sphere S.

4.5 Joint moments and linearity of conditional expectations.

Recall that if X is a one dimensional stable variable with index α, then $E|X|^p < \infty$ for all $0 < p < \alpha$. The moment of order equal to α need not be finite as the Cauchy law where $\alpha = 1$ shows. Thus when we deal with stable vectors, we must at least assume that $\alpha > 1$ for the moments to exist in general. By using Holder's inequality, this not only assures the finiteness of the first moment of *every* component, it also implies every product moment of combined order $p < \alpha$ is finite. To be precise, if X is stable with index α, then

$$\sum_{i=1}^{n} p_i < \alpha, \quad \Rightarrow E\prod_{i=1}^{n} |X_i|^{p_i} < \infty.$$

The converse is false in general. However if $X = (X_1,\ X_2)$ is a stable vector with only *two* coordinates, and with index $\alpha < 2$ then, the converse is indeed true.

What happens to regression functions for stable vectors? In particular, if $X = (X_1, \ldots, X_n)$ is stable, is $E(X_1|X_2, \ldots, X_n)$ linear in $X_2, ...X_n$?

Again, the answer is yes, if we have a two vector: if $X = (X_1,\ X_2)$ is stable with index $1 < \alpha < 2$, then $E(X_2|X_1) = cX_1$ for some constant c. If $X = (X_1, \ldots, X_n)$, $n > 2$, then in general it is *not* true $E(X_1|X_2, \ldots, X_n)$ is linear in $(X_2, \ldots, X_n)$.

There are several conditions under which this linearity can be claimed. We give one which is related to the spectral measure Γ.

If X is strictly stable with index $1 < \alpha < 2$, and spectral measure Γ, then, $E(X_n|X_1, \ldots, X_{n-1} = \sum_{i=1}^{n-1} a_i X_i$ if and only if

$$\forall\ \underset{\sim}{r},\quad \int_S (x_n - \sum_{i=1}^{n-1} a_i x_i)(\sum_{i=1}^{n-1} r_i x_i)^{\alpha-1} d\Gamma\ (x) = 0$$

5. Simulation of id laws

To understand the behavior of different statistical procedures where id laws are involved, it is important to be able to simulate id and stable laws. We shall concentrate here on simulation of id laws in general. The simulation of stable laws is a significantly more specialized task and will not be discussed here. The interested reader may consult Adler et. al. (1998) for material on simulation of stable laws. We discuss two approaches.

Approach 1: via Poisson processes. Bondesson (1982) noticed an interesting connection between Poisson processes and id laws as follows.

Let $Z(u)$, $u > 0$ be a family of non-negative independent random variables. Let T_i, $i = 1, 2\ldots$ be the points (in increasing order) in an independent Poisson point process of rate λ on $(0,\ \infty)$. We set

$$X = \sum_{i=1}^{\infty} Z(T_i).$$

Note that if we define

$$X(t) = \sum_{T_i' \leq t} Z(t - T_i')$$

then it is a *shot noise* process and $X = X(0)$. So the distribution of X is called a shot noise distribution.

Set

$$X_T = \sum_{T_i \leq T} Z(T_i)$$

where T is a truncation point. Its Laplace transform (LT) is given by

$$E[\exp\{-sX_T\}] = \exp\left\{\lambda \int_{(0,\ T)} (\psi(s,u) - 1) du\right\}$$

where

$$\psi(s,u) = E[\exp\{-sZ(u)\}]. \tag{6}$$

This can be proved in at least two ways. One proof is based on a conditioning on the number of points in $(0,\ T]$, and the other, more stringent one, is based on an approximation (in distribution) of X_T as $\sum_{k=1}^{[T_n]} Z(k/n)I_k^n$, where I_k^n is one or zero depending on whether or not there are any points in $(k-1/n,\ k/n]$. It follows that X has the LT

$$\phi(s) = \exp\left\{\lambda \int_{(0,\ \infty)} (\psi(s,u)-1)du\right\}. \tag{7}$$

We assume that $X < \infty$ almost surely. (Otherwsie $X = \infty$ almost surely: this follows from the zero one law and the fact that the process $X_T,\ T > 0$ has independent increments.)

It is obvious that X is id.

Let $Z(u)$ have the distribution function $H(y,u)$. Changing the order of integration, we may rewrite (7) as

$$\phi(s) = \exp\left\{\lambda \int_{[0,\ \infty)} (e^{-sy}-1)\left(\int_{(0,\ \infty)} H(dy,u)du\right)\right\} \tag{8}$$

Of course the measure $H(dy,u)$ must as a function of u satisfying certain regularity conditions.

Now we consider simulation from an id distribution F with Levy measure $\Lambda(dy)$. Suppose we can find a simple family of distribution functions $H(y,u)$ on $[0,\ \infty)$ and a λ such that on $(0,\ \infty)$,

$$\lambda \int_{(0,\ \infty)} H(dy,u)du = N(dy)$$

or equivalently, for $x > 0$,

$$\lambda \int_{(0,\ \infty)} \bar{H}(x,u)du = \int_{(x,\ \infty)} N(dy) = \bar{N}(x)$$

where $\bar{H} = 1 - H$. Then simulate points T_i in a Poisson (λ) process by for example adding independent exponential randon numbers and after that, values $Z(T_i)$ from the distribution functions $H(x,\ T_i)$ and set $X = \sum_{i=1}^{\infty} Z(T_i)$. Then X has the desired distribution. If the sum converges rapidly, only a few terms are needed to get a good approximate value of X.

Bondesson (1982) showed how different classes of H lead to different id distributions such as, the generalized convolutions of mixtures of exponentials (class $\mathcal{T}_2$ of Bondesson (1981)), generalized gamma convolutions (Thorin (1977, 1978)) and the generalized negative binomial convolutions (Bondesson (1979)).

Approach 2: via structural theorem(characterization #4). Recall characterization #4 given earlier for id laws as limit of sums of independent Poisson type random variables. We state this fact again here in the form of a theorem, commonly known as the *structural theorem.* A proof may be found in Loeve (1960, page 298).

Theorem: A characteristic function ψ is id if and only if it is the limit of sequences of products of Poisson types. That is there exists a_{nk} and b_{nk} such that

$$\psi(t) = \lim_{n\to\infty} (\exp[\sum_{k=1}^{n} ita_{nk} + \lambda_{nk}\{\exp(-itb_{nk}) - 1\})$$

The algorithm of Damien, Laud and Smith (1995) to generate an observation from a given id law with characteristic function ψ proceeds as follows:

Let $\wedge$ be the appropriately defined (finite) Levy-Khintchine measure associated with ψ.

Let $\wedge_1, \ldots, \wedge_n$ be i.i.d. from the distribution $\frac{1}{k} d\wedge(x)$ where $k = \int_{-\infty}^{\infty} d\wedge(x)$.

Let $Y_i \sim$ Poi $(\frac{k(1+\wedge_i^2)}{n\wedge_i^2})$, $i = 1, \ldots, n$

Let $X_n = \sum_{i=1}^{n} (\wedge_i Y_i - \frac{k}{n\wedge_i})$. Then $\psi_{X_n}(t) \longrightarrow \psi(t)\ \forall\ t$, as $n \to \infty$.

In particular, they use this algorithm to generate observations from several stable distributions and study the accuracy via the Kolmogorov-Smirnov metric.

This has interesting applications in Bayesian nonparametrics. Consider the problem of estimating an unknown cdf F on $[0, \infty)$ based on n iid observations (possibly censored) from F. This requires putting a prior distribution on the space of distribution functions $\mathcal{F}$. Viewing F as a stochastic process, let $F(t) = 1 - \exp(-Y_t)$ where $\{Y_t\}$ is a Levy process, that is, a process having $Y(t+s) - Y(t) > 0$; independent $\forall\ s,\ t > 0$). The posterior distribution is also a Levy process. See Ferguson and Phadia for details. The increments of this process, when the jumps are removed, are id. Using the above approach, these continuous increments can be simulated. The jump components are independent and hence simulating the increments corresponding to these jumps is standard. Combining these two simulations, the total increments of the process are simulated. This implies that a complete Bayesian analysis of the posterior distribution is possible. In particular, the authors show how to implement the idea for estimating the survival function using the three priors, gamma process, Dirichlet process and the simple homogeneous process.

Apparently, no results are known regarding the rate of convergence of the generated samples, but the simulation results of the authors are quite promising.

Related papers in Bayesian nonparametrics where particular Levy processes have been used are (i) Hjort (1990) who uses beta processes, and (ii) Ramgopal and Smith (1993) who use extended gamma processes.

6. Stable laws in inference

As we have seen earlier, no id law can have tails thinner than the normal tail. However, the tails of an id law can be quite heavy. As data from a steadily increasing number of fields have exhibited heavy tailed behavior, the importance of id laws in statistical modelling and inference has grown.

We have very briefly mentioned the use of id processes in Bayesian inference in the previous section. However, since the class of id laws consists of the weak limits of triangular sums, it is a huge class and is not convenient for most statistical modelling and inference problems.

On the other hand, any stable law is obtained as the weak limit of sums of iid random variables. Thus it serves as a very natural model in situations where aggregation is involved. This explains the importance of the normal distribution when the observations have finite second moments. But this leaves out the distributions with heavy tails.

As we have seen in section 3.3, at least one of the tails of a stable law decreases as the αth power. This offers flexibility in modelling heavy tailed phenomena by stable laws with an appropriate choice of α, $0 < \alpha < 2$. Instances where the stable model holds exactly are not very frequent. The earliest known example came before the stable laws were formally introduced by P. Levy.

Example 24. As early as 1919, before the concept of stable laws was introduced by Paul Levy, Holtsmark found that under certain natural assumptions, the random fluctuations of the gravitational field of stars in space has a probability density whose cf is given by $\exp\{-\lambda|t|^{3/2}\}$, $t \in \mathbb{R}^3$ where λ is a positive constant determined by certain physical characteristics. This is a three dimensional spherically symmetric stable law with $\alpha = 3/2$ and is known as the *Holtsmark distribution.*

Since we cannot hope to have exact stability of the observations we must look for approximate stability. This leads to the concept of domain of attraction.

6.1 Domain of attraction.

Definition 5. A distribution F is said to belong to the *domain of attraction* of a stable law with index α if there exists real sequences $\{a_n > 0\}$ and $\{b_n\}$ such that if $X_1, \ldots X_n \ldots$ are iid with distribution F then $b_n^{-1}(X_1 + \ldots + X_n - a_n)$ converges in distribution to this stable law. We write $F \in \mathcal{D}(\alpha)$.

Example 25. Any stable distribution is trivially in its own domain of attraction. All distributions with finite second moments are in the domain of attraction of the normal law.

Plenty of distributions with infinite second moments are also in the domain of the normal law. It will be easier to provide such examples after we give the criteria for checking whether a distribution belongs to $\mathcal{D}(\alpha)$.

There are two such simple but powerful criteria. To state these, recall that a function $L(\cdot)$ is said to be *slowly varying* if for every $x > 0$, $L(tx)/L(t) \to 1$ as $t \to \infty$. Below, L is any such slowly varying function.

Criterion 1. A distribution F belongs to the domain of attraction of a stable law

(i) with index $0 < \alpha < 2$, if and only if there exists $0 \leq p \leq 1$ such that,

$$\lim_{x \to \infty} \frac{1 - F(x)}{1 - F(x) + F(-x)} = p \tag{9}$$

and

$$1 - F(x) + F(-x) \sim \frac{2 - \alpha}{\alpha} x^{-\alpha} L(x) \text{ as } x \to \infty \tag{10}$$

(ii) with index $\alpha = 2$ (normal law), if and only if

$$\int_{-x}^{x} y^2 dF(y) \sim L(x) \text{ as } x \to \infty. \tag{11}$$

Example 26. Consider the Pareto law discussed in Example 10 which has the density $f(x) = \frac{\alpha}{\mu}(\frac{\mu}{x+\mu})^{\alpha+1}$, $x > 0$. It is easy to see that it belongs to $\mathcal{D}(\alpha)$.

Example 27. By using the criterion above, it is easy to construct examples of distributions F whose second moments are infinite but which belong to the domain of attraction of the normal law. For instance, the distribution F with density $f(x) = 2|x|^{-3} \log x$ for $|x| \geq 1$ has infinite second moment and belongs to the domain of attraction of the normal law. The t distribution with one degree of freedom is the Cauchy law and so is stable. The t distribution with degrees of freedom three or more has finite second moment and hence is in the the domain of attraction of the normal law. The t distribution with *two* degrees of freedom has the density $f(x) = c(1 + x^2)^{-3/2}$ where c is a constant. So it does not have finite second moment. However, it is easy to check that Criterion 1 (ii) is satisfied with $L(x) = \log x$. Hence the t distribution with two degrees of freedom belongs to the domain of attraction of the normal law.

Example 28. In the definition of domain of attraction, we used *sums* of variables. If we use other composition operations, we obtain other notions of stability. While we do not wish to present all such notions, we wish to discuss one such alternate notion of stability, obtained by taking *maximums*. Suppose that $X_1, \ldots, X_n$ are iid F. Let $M_n = \max\{X_1, \ldots, X_n\}$. Suppose that $\{a_n > 0\}$ and $\{b_n \in \mathbb{R}\}$ are sequences such that $a_n^{-1}(M_n - b_n)$ converges in distribution to some distribution G. We write $F \in \max\mathcal{D}(G)$. The class of limit distributions obtained in this way is called the class of *extreme value distributions* or *max stable laws.* A parametric description of this class is given in section 6.4. One subclass of this class consists of the *Frechet distributions* defined as:

$$\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}, \; x > 0, \; \text{where } \alpha > 0.$$

It is known from the extreme value theory that $F \in \max\mathcal{D}(\Phi_\alpha)$ if and only if $1 - F(x) = x^{-\alpha}L(x)$. Note that this condition is a *part* of the condition for F to belong to $\mathcal{D}(\alpha)$. We discuss the use of connection between stable laws and max stable laws in statistical estimation in section 6.4.

We now present the second domain of attraction criterion. An application to the problem of estimation of α may be found in section 6.3.

Criterion 2. A distribution F belongs to the domain of attraction of a stable law with index $0 < \alpha < 2$, if and only if

$$\frac{x^2[1 - F(x) - F(-x)]}{\int_{-x}^{x} y^2 dF(y)} \sim \frac{2-\alpha}{\alpha} \text{ as } x \to \infty. \tag{12}$$

6.2 Estimation of α, preliminaries.

The normal law has a rapidly decreasing tail and corresponds to $\alpha = 2$. For modelling heavy tailed phenomena, we restrict our discussion to the class of stable laws with index $0 < \alpha < 2$. This leads to the following basic question:

Question: Suppose we have iid observations from a distribution $F \in \mathcal{D}(\alpha)$. How does one estimate the parameter α?

Note that even if we assume that F itself is stable, the problem is still not easy. As mentioned earlier, except for the three special distributions normal, Cauchy and Levy, no closed form expressions are known for the density of stable laws. This makes the problem of estimating α quite difficult. Possible approaches to the estimation problem are already offered indirectly in the discussion of section 6.1:

(i) Example 27 suggests that the extreme order statistics have a role to play.

(ii) Criterion 1 suggests how the sample versions of F, namely the *empirical distribution* F_n may be used to obtain estimates of α. Likewise, Criterion 2 also suggests estimates for α.

(iii) The cf of stable laws is available in a closed form, Thus the use of *empirical characteristic function* offers another possible approach, at least when F is exactly stable.

Example 29. Consider the *one* parameter Pareto distribution with parameter α whose cdf is given by

$$1 - F(x) = x^{-\alpha}, \quad x > 1. \tag{13}$$

Assume that $X_1, \ldots, X_n$ are iid observations from this Pareto law. Since the distribution and the density in this case are explicitly known, we can use the *method of maximum* *likelihood* to estimate α. By writing down the joint density of $X_1, \ldots, X_n$, it is easily seen that the *maximum likelihood estimator* of $\gamma = \alpha^{-1}$ is given by

$$\hat{\gamma}_n = n^{-1} \sum_{i=1}^{n} \log X_i \ = n^{-1} \sum_{i=1}^{n} \log X_{(i)} \tag{14}$$

Above, $X_{(1)} < \ldots < X_{(n)}$ are the *order statistics* of $X_1 \ldots X_n$. We shall use this notation in our subsequent discussion also.

The nice thing about this estimate is that it involves the random variables through their logarithms which have finite second moments. Indeed,

$$E(\log X_1) = \gamma \quad \text{and } Var(\log X_1) = \gamma^2. \tag{15}$$

By using the central limit theorem, we thus have

$$n^{1/2}(\hat{\gamma}_n - \gamma) \stackrel{\mathcal{D}}{\Rightarrow} N(0,\ \gamma^2). \tag{16}$$

Suppose now that $F \in \mathcal{D}(\alpha)$. Suppose that the right tail is nontrivial so that equation (9) is satisfied with $p > 0$. Then Criterion 1 implies that the right tail behaves like the Pareto tail (13) in Example 28, except for a slowly changing function. This feature is the basis of many estimators of α in the literature. In the next few subsections we shall describe some of the estimators of α. For some comparisons of these estimators based on simulations, see Pictet et. al. (1998). The general recommendation is that Hill's estimator, discussed in section 6.3, is the best to use.

6.3 The Hill estimator.

Assume that $F \in \mathcal{D}(\alpha)$ is such that

$$1 - F(x) = x^{-\alpha} L(x), \quad \text{as} \quad x \to \infty \tag{17}$$

where $L(\cdot)$ is a *slowly varying function.* Note that this implies that *if* $X_{(n-k)}$ *is large,* then the following approximate relation holds:

$$\frac{1 - F(x\ X_{(n-k)})}{1 - F(X_{(n-k)})} \approx x^{-\alpha}. \tag{18}$$

Conditional on $X_{(n-k)}$, $\left(\frac{X_{(n)}}{X_{(n-k)}}, \ldots, \frac{X_{(n-k+1)}}{X_{(n-k)}}\right)$ is distributed as the order statistics from a sample of size k from the distribution with tail

$$\frac{1 - F(x\ X_{(n-k)})}{1 - F(X_{(n-k)})}, \quad x \geq 1$$

which, as (17) holds, is approximately the Pareto tail. Thus going back to the estimate introduced in the special case of the Pareto, it appears intuitively justified to use the above ratios, for some large value of k, in the same way as all the observations were used in defining (14) in the exact Pareto case. This leads to the famous Hill's estimator (Hill (1975)): choose $k < n$ large in some appropriate way. Then the *Hill estimate* of $\gamma = \alpha^{-1}$ on the basis of n iid observations from the distribution F satisfying (17) is defined as

$$\hat{\gamma}_{k,n} = k^{-1} \sum_{i=n-k+1}^{n} \log \frac{X_{(i)}}{X_{(n-k)}}. \tag{19}$$

The Hill estimate uses only the *upper* $(k+1)$ ordered statistics of the sample and ignores the rest of the sample. The uneasy aspect is the dependence on the choice of k. We shall address this issue below. But first let us see a result which guarantees that this method works, at least asymptotically.

Consistency of the Hill estimator. Suppose that $n \to \infty$ so that we have a sample size which increases indefinitely. Let $k = k_n$ be such that $k \to \infty$ but $k/n \to 0$. This means that we use a very large proportion of the ordered statistics but leave out a significant fraction. It turns out that this guarantees that (Mason (1982))

$$\hat{\gamma}_{k,n} \xrightarrow{P} \gamma. \tag{20}$$

Note that no additional assumptions on F are required for the above result. So the Hill estimator is indeed *consistent* under minimal assumptions.

In practice, one has to deal with data which are not iid. Extensions of the above consistency to situations where $\{X_i\}$ is a dependent sequence may be found in Rootzen, Leadbetter and de Haan (1990), Hsing (1991), and Resnick and Stărică (1998).

Asymptotic distribution and confidence interval. In applications, one is not satisfied with a *point* estimate and a consistent *interval* estimate is more comforting. This requires establishing a nondegenerate (asymptotic) distribution of the estimator with an appropriate norming and centering. Unfortunately, the class of all F which are in the domain of a stable law with index α is still too large and such a result is *not* available. However, under suitable restrictions on F the same limit law (16) as in the exact Pareto case holds. This result is actually true under several different sets of sufficient conditions. The reader may consult de Haan and Resnick (1998) and the references contained there for more details. *Under suitable conditions on* k and F,

$$k^{1/2}(\hat{\gamma}_{k,n} - \gamma) \xrightarrow{\mathcal{D}} N(0,\ \gamma^2). \tag{21}$$

It is assuring that the limit distribution is *normal* and the limiting variance involves F only through γ. This makes setting up an approximate confidence interval for γ easy. Fix a confidence coefficient $1-\beta$ and let $\Phi^{-1}(\beta/2)$ be the *upper* $\beta/2$ percentile of the standard normal distribution. Then a $100(1-\beta)\%$ asymptotically correct confidence interval for γ is given by:

$$I_n = [\hat{\gamma}_{k,n}\{1+k^{-1/2}\Phi^{-1}(\beta/2)\},\ \hat{\gamma}_{k,n}\{1-k^{-1/2}\Phi^{-1}(\beta/2)\}]. \tag{22}$$

The consistency result (20) and the asymptotic normality result (21) together imply that $P\{\hat{\gamma}_{k,n} \in I_n\} \to 1-\beta$ under the conditions alluded to. The equivalent statement for the estimate of α is of course obtained by taking the interval with the end points as the inverses of the end points of I_n.

Choice of k: the Hill plot. The consistency and asymptotic normality property of the Hill estimator depends on $k = k_n$ going to infinity at an appropriate rate. In practice, given a sample of size n, one has to decide on the value of k to use. One approach is to use the *Hill plot.* This is simply a plot of the estimator $(\hat{\gamma}_{k,n})^{-1}$ against k. On this plot we look for a range of values of k where the plot is flat. This gives a range of possible values of k which can be used to calculate the estimate. Empirically, it has been seen that the estimator is quite insensitive to the eventual choice of k in the chosen range. For more information, see Drees et. al. (2000). This article also carries information on various refinements of the Hill plot.

Bias of the Hill estimator in small samples. Since the Hill estimator is based on an approximation of the tail of F, it is natural for it to have some *bias* in finite samples. The amount of the bias is determined by the finer behavior of the tail of F. One possibility in studying the bias is to work with specified subclasses of F. Here is one such result. Consider the class of $F \in \mathcal{D}(\alpha)$ which satisfy for some $a > 0$ and $\beta > 0$,

$$1 - F(x) = ax^{-\alpha}[1 + bx^{-\beta} + o(x^{-\beta})] \tag{23}$$

Then if $k = k_n \to \infty$, $k/n \to 0$, the asymptotic bias B of the Hill estimator is given by:

$$B = -\frac{\beta b}{\alpha(\alpha+\beta)} a^{-\frac{\beta}{\alpha}} (\frac{k}{n})^{\frac{\beta}{\alpha}} \{1 + o(1)\} \tag{24}$$

An asymptotic expression for the variance can also be derived:

$$Var(\hat{\gamma}_{k,n}) = [\frac{\beta^2 b^2}{\alpha^2(\alpha+\beta)^2} a^{-\frac{2\beta}{\alpha}} (\frac{k}{n})^{\frac{2\beta}{\alpha}} + \frac{1}{\alpha^2 k}] + o(1). \tag{25}$$

See Goldie and Smith (1987), Hall and Welsh (1985) and Pictet et. al. (1998) for bias expressions in various situations and recommendations for the choice of k.

Unsatisfactory behavior near $\alpha = 2$. While the Hill estimator is one of the best and popular methods, its unsatisfactory performance is documented in the literature when α is close to 2. A possible explanation is that while the tail of a stable law with index $\alpha < 2$ is like $x^{-\alpha}$, the tail of the normal law ($\alpha = 2$) is exponentially decreasing. Further, a few upper order statistics cannot be expected to yield good estimators for "near normal" laws.

6.3 de Haan and Pereira's estimator.

de Haan and Pereira (1999) focussed on the situation where α may be close to 2. Let $\beta = \frac{2-\alpha}{\alpha}$. Note that if α is close to 2 then β is close to zero. In this situation, it appears to be reasonable to consider Criterion 2 and start with (12) to build an estimator. Consideration of the sample analogue of Criterion 2 leads to their estimator.

So suppose we have iid observations on $F \in \mathcal{D}(\alpha)$. Let the order statistics of $|X_i|$, $1 \le i \le n$ be denoted by $|X|_{(1)} \le \ldots \le |X|_{(n)}$. Let $\mathcal{G}_n$ be the empirical distribution of $\{|X_i|,\ 1 \le i \le n\}$. Motivated by (12), we may choose a $k = k_n \to \infty$ and define the estimator $\hat{\beta}_n$ of $\beta = \frac{2-\alpha}{\alpha}$ as

$$\hat{\beta}_n = \frac{k|X|^2_{(n-k)}}{\sum\limits_{i=1}^{n-k} |X|^2_{(i)}}. \tag{26}$$

It may be noted that this estimate uses the $(n-k)$ *lower* order statistics of the absolute values. The estimate of β is easily transformed into an estimate $\hat{\alpha}_n$ of α as $\hat{\alpha}_n = 2(1+\hat{\beta}_n)^{-1}$.

Under various assumptions on $\{k_n\}$, and F, the consistency and asymptotic normality of $\hat{\beta}_n$ hold. However, the norming is not as simple as the one in the Hill estimator. As with the Hill estimator, in our statement, we shall leave out the exact assumptions required. For details of the conditions required, see de Haan and Pereira (1999). To state the asymptotic normality of $\hat{\beta}_n$, let

$$N_n = \frac{1}{2}\sum_{i=0}^{k-1}\{\log|X|^2_{(n-i)} - \log|X|^2_{(n-k)}\}$$

$$\beta_n = \frac{k}{n}\{\mathcal{G}^{-1}(1-k/n)\}^2 \Big/ \int_0^{1-\frac{k}{n}} (\mathcal{G}^{-1}(t))^2 dt$$

where $\mathcal{G}^{-1}$ is the inverse of $\mathcal{G}(x) = F(x) - F(-x)$. Then *under appropriate conditions,*

$$\frac{k^{1/2}}{N_n}\left(\frac{\hat{\beta}_n}{\beta_n} - 1\right) \xrightarrow{\mathcal{D}} N(0,\ (2\beta+1)^{-1}) \tag{27}$$

6.4 A moment estimator.

In Example 27 we have seen a parametric subclass of the max-stable laws. It turns out that the entire collection of extreme value distributions can also be parametrized. The approach to estimating this parameter leads to an estimate for the stability index α as well.

A distribution is an extreme value distribution if and only if up to a scale and location shift, it is of the form:

$$G_\gamma(x) = \exp(-(1+\gamma x)^{-\gamma^{-1}}), \quad \gamma \in \mathbb{R}, \; x > 0. \tag{28}$$

The case of $\gamma = 0$ is interpreted as

$$G_0(x) = \exp(-e^{-x}). \tag{29}$$

The parameter γ may be called the *extreme value index* of the distribution. For $\gamma > 0$, let $\alpha = \gamma^{-1}$. Then this parametrization is consistent with the parametrization of the stable class. That is, $F \in \mathcal{D}(\alpha)$ for some $0 < \alpha < 2$, if and only if $F \in \max \mathcal{D}(G_\gamma)$. Now consider the problem of estimating γ when it is known that $F \in \max \mathcal{D}(G_\gamma)$. Suppose that $\gamma > 0$. Dekkers, Einmahl and de Haan (1989) considered the problem of estimation of γ and one of the estimators they consider is obtained by a moment approach. For $r = 1, 2$, let

$$H_{k,n}^{(r)} = k^{-1} \sum_{i=n-k+1}^{n} \left(\log \frac{X_{(i)}}{X_{(n-k)}} \right)^r. \tag{30}$$

Hence, $H_{k,n}^{(1)}$ is Hill's estimator. Define the estimator $\hat{\gamma}_n$ of γ as

$$\hat{\gamma}_n = H_{k,n}^{(1)} + 1 - \frac{1/2}{1 - (H_{k,n}^{(1)})^2 / H_{k,n}^{(2)}}. \tag{31}$$

Note that the estimator $\hat{\gamma}_n$ is an estimator of the extreme value index γ and is defined even if F does not belong to $\mathcal{D}(\alpha)$. That is, it is an estimator of γ irrespective of whether $\alpha = 1/\gamma$ is in the interval (0, 2). This is an important aspect: suppose we do not know whether F has heavy tails. Then we will be wary of using the Hill estimator since it is specially geared for the heavy tailed situation. We can then consider using the current estimator.

The estimator is consistent for all values of γ: if $F \in \max \mathcal{D}(\gamma)$, $k \to \infty$ and $k/n \to 0$, then

$$\hat{\gamma}_n \xrightarrow{P} \gamma. \tag{33}$$

So then why use Hill's estimator at all? This is reflected in the asymptotic distribution of the estimator. As before we skip the precise conditions, which can be seen in Dekkers et. al. (1990). Under suitable conditions,

$$k^{1/2}(\hat{\gamma}_n - \gamma) \xrightarrow{D} N(0, 1 + \gamma^2) \tag{34}$$

Recall that the asymptotic variance of the Hill estimator is γ^2 and so, the current estimator has a larger asymptotic variance than the Hill estimator.

6.5 Other estimators.

There are many other estimators that are available in the literature. We will not go into a detailed description of these. Here are some of the more well known ones:

1. Pickands estimator: This is a very quick and easy estimator proposed in Pickands (1975). It involves calculating the 25%, 50% and 75% quantiles. See Dekkers and de Haan (1989) for its strong consistency and asymptotic normality under appropriate conditions. The estimator is defined as

$$\hat{\gamma}_n^P = (\log 2)^{-1} \log \frac{X_{(k)} - X_{(2k)}}{X_{(2k)} - X_{(4k)}}. \tag{35}$$

2. de Haan-Resnick estimator: This estimator is given in de Haan and Resnick (1980) and involves only the maximum and *one* other extreme order statistics. It is thus a simplified version of the Hill estimator.

$$\hat{\gamma}_n^R = \frac{\log X_{(1)} - \log X_{(k)}}{\log k}. \tag{36}$$

3. The CD plot estimator. The log-log *complementary distribution* (CD) plot estimator also has its genesis in the Pareto expression

$$1 - F(x) \sim x^{-\alpha} \text{ as } x \to \infty. \tag{37}$$

This implies that $\log(1 - F(x))$ and x are linearly related *for large* x with slope $-\alpha$. In practice, we plot $\log(1 - F_n(x))$ against x and choose a *large* x_0 beyond which the plot looks linear. Estimate the (negative) slope by fitting a straight line (with equally spaced chosen points on the X-axis) and the negative of the slope is the estimate for α.

Remark 18. Even though in practice observations can rarely be assumed to be exactly stable, it is illuminating to consider such a situation and investigate how the different parameters (α, β and b) in the corresponding cf representation can be estimated. These estimators can also serve as preliminary estimators in more complicated procedures which involve observations which are not exactly stable. The *McCulloch estimator* is the simplest among these and is designed for the situation when the observations are from a stable law with $\alpha \in [0.6,\ 2]$. The main virtue of the estimator is its simplicity of calculation. It may be termed as the *method of five quantiles* and is known to perform remarkably well in practice.

Suppose that F is stable with the cf given in section 3.3. Let F_p denote the pth quantile of F. Let

$$\Phi_1(\alpha,\ \beta) = \frac{F_{0.95} - F_{0.05}}{F_{0.75} - F_{0.25}} \text{ and } \Phi_2(\alpha,\ \beta) = \frac{F_{0.95} + F_{0.05} - 2F_{0.50}}{F_{0.95} - F_{0.05}}. \qquad (38)$$

It turns out that Φ_1 is monotonic in α and Φ_2 is monotonic in β (for fixed α) and so we can invert these functions to get

$$\alpha = \Psi_1(\Phi_1,\ \Phi_2) \text{ and } \beta = \Psi_2(\Phi_1,\ \Phi_2). \qquad (39)$$

McCulloch (1986) tabulated these values for various values of Φ_1 and Φ_2.

To form the estimators of α and β, first estimate the five quantiles above by the respective sample quantiles. Use these to obtain estimates of Φ_1 and Φ_2. Then use McCulloch's tables to obtain the estimates of the two parameters.

Another common approach is to use the representation of the characteristic functions of stable distributions. The corresponding sample cf is used to build up these estimators. It will take a lengthy treatment to do justice to them. The reader may consult Kogon and Williams (1998) and the references contained in that paper for material on this topic.

7. Applications

Stable and infinitely divisible distributions have found the greatest aplications in finance and economics. There have been other applications as well in problems involving heavy tails; see the recent book by Uchaikin and Zolotarev (1999). Here we will mention a few applications in the areas of finance and economics.

Benoit Mandelbrot made the first attempt to use stable distributions for modeling stock returns by questioning the use of normal distributions for that purpose; see Mandelbrot (1963). Use of stable laws for analyzing stock returns is also made in Officer (1972).

Applications in capital asset pricing are discussed in Gamrowski and Rachev (1995), and in a very nice review article by McCulloch (1996). Stable laws have also been used in option pricing and for modeling foreign exchange rates; see McCulloch (1996) for comprehensive review of the models.

The finance and economics literature also contain methods for estimation of stable law parameters, and this development has been partially independent of the probability and statistics literature. Methods of parameter estimation are discussed in Arad (1980) in the context of stock returns, and in Liu and Brorsen (1995) in the context of modeling foreign exchange rates, in particular.

References

Adler, R. J., Feldman, R. E. and Taqqu, M. S. (Editors) (1998). *A practical guide to heavy tails: statistical techniques for analyzing heavy tailed distributions.* Birkhäuser, Boston.

Adler, R.and Taqqu, M. (1997). Structure of stationary stable processes. In *A practical guide to heavy tails.* Birkhäuser, Boston.

Akgiray, V. and Lamoureux, C. G. (1989). Estimation of stable-law parameters: a comparative study. *J. Business and Statistics*, **7**, 1, 85–93.

Arad, R. W. (1980). Parameter estimation for symmetric stable distributions. *Int. Eco. Rev.* 21, 209–220.

Arnold, Barry C. and Wesolowski, Jacek (1997). Distributions with Gaussian conditional structure. Stochastic Processes and Functional Analysis, *Lecture Notes in Pure and Applied Mathematics*, 186, 45–59. Dekker, New York.

Ash, Robert B. (1972). *Real Analysis and Probability.* Academic Press, New York.

Athreya, K. B. and Lahiri, S. N. (1998). Inference for heavy tailed distributions. *J. Statist. Planning and Inf.* **66**, 61–75.

Barndorff-Nielsen, O., Mikosch, T. and Resnick, S.I. (Eds). (2001). *Levy Processes:theory and applications.* Birkhauser, New York.

Bartels, R. (1978). Generating non-normal stable variates using limit theorem properties. *J. Statist. Comput. Simul.* **7**, 199–212.

Bartels, R. (1980/81). Truncation bounds for infinite expansions for the stable distributions. *J. Statist. Comput. Simulation* **12**, 3–4, 293–302.

Basawa, I. V. and Brockwell, P. J. (1978). Inference for gamma and stable processes. *Biometrika* **65**, 1, 129–133.

Basawa, I. V. and Brockwell, P. J. (1980). A note on estimation for gamma and stable processes. *Biometrika* **67**, 1, 234–236.

Basu, Sujit K. and Sen, Pranab Kumar (1985). Asymptotically efficient estimator for the index of a stable distribution. *Proceedings of the seventh conference on probability theory* (Braşov, 1982) 33–39, VNU Sci. Press, Utrecht, 1985.

Beirlants, J. and Teugels, J. K. (1986). Asymptotics of Hill's estimator. *Theory Probab. Appl.* **31** 463–469.

Berman, S. M. (1986). The supremum of a process with stationary independent and symmetric increments. *Stoch. Proc. Appl.* 23, 2, 281–290.

Billingsley, Patrick (1995). *Probability and Measure.* Third Edition. John Wiley & Sons, Inc. New York.

Blum, J. R. and Rosenblatt, M. (1959). On the structure of infinitely divisible distributions. *Pac. J. Math.*, **9**, 1–7.

Bondesson, Lennart (1992). *Generalized gamma convolutions and related classes of distribution and densities.* Lecture Notes in Statistics, 76. Springer - Verlag, New York.

Bondesson, L. (1978). On infinite divisibility of powers of a gamma variable. *Scand. Actuar. J.* **1**, 48–61.

Bondesson, L. (1979). On the infinite divisibility of products of powers of gamma variables. *Z. Wahrsch. Verw. Gebiete* **49**, 2, 171–175.

Bondesson, L. (1979). A general result on infinite divisibility. *Ann. Probab.* **7**, 6, 965–979.

Bondesson, L. (1981). Classes of infinitely divisible distributions and densities. *Z. Wahrsch. Verw. Gebiete* **57**, 1, 39–71.

Bondesson, L. (1982a). Correction and Addendum to : "Classes of infinitely divisible distributions and densities. *Z. Wahrsch. Verw. Gebiete* **59**, 2, 227.

Bondesson, L. (1982b). On simulation from infinitely divisible distributions. *Adv. Appl. Probab.* **14**, 855–869.

Bondesson, L. (1987). On the infinite divisibility of the half-Cauchy and other decreasing densities and probability functions on the nonnegative line. *Scand. Actuar. J.* **3–4**, 225–247.

Bondesson, L. (1988). A remarkable property of generalized gamma convolutions. *Probab. Theory Related Fields* **78**, 3, 321–333.

Bondesson, L. (1990). Generalized gamma convolutions and complete monotonicity. *Probab. Theory Related Fields* **85**, 2, 181–194.

Bondesson, L. (1992). *Generalized gamma convolutions and related classes of distributions and densities.* Lecture Notes in Statistics 76. Springer Verlag.

Bondesson, L., Kristiansen, K. and Steutel, F. W. (1996). Infinite divisibility of random variables and their integer parts. *Statist. Probab. Lett.* **28**, 3, 271–278.

Breiman, Leo (1992). *Probability.* Society for Industrial and Applied Mathematics, Philadelphia.

Brockwell, P. J. and Brown, B. M. (1981). High-efficiency estimation for the positive stable laws. *J. Amer. Stat. Assoc.*, 76, 375, 626–631.

Brockwell, P. J. and Liu, Jian (1992). Estimating the noise parameters from observations of a linear process with stable innovations. **33**, 2, 175–186.

Brockwell, P. J., and Mitchell, Heather (1998). Linear prediction for a class of multivariate stable process. Special Issue in Honor of Marcel F. Neuts. *Comm. Statist. Stochastic Models* **14**, 1–2, 297–310.

Brown, L. and Rinott, Y. (1988). Inequalities for multivariate infinitely divisible processes. *Ann. Probab.*, 2, 642–657.

Brown, T. C., Feigin, Paul, D. and Pallant, Diana L. (1996). Estimation for a class of positive nonlinear time series models. *Stochastic Process. Appl.* **63**, 2, 139–152.

Buckle, D. J. (1995). Bayesian inference for stable distributions. *J. Amer. Statist. Assoc.* **90** 430, 605–613.

Byczkowski, T., Nolan, J. P. and Rajput, B. (1993). Approximation of multidimensional stable densities. *J. Mult. Anal.*, **46**, 13–31.

Carnal, H. and Dozzi, M. (1989). On a decomposition problem for multivariate probability measures. *J. Mult. Anal.*, **2**, 165–177.

Chamber, J. M., Mallows, C. M. and Stuck, B. W. (1976). A method for simulating stable random variables. *J. Amer. Statist. Assoc.* **71**, 354, 340–344.

Chang, Derek K. (1989). On infinitely divisible discrete distributions. *Utilitas Math.* **36**, 215–217.

Chow, Yuan Shih and Teicher, Henry (1997). *Probability Theory, Independence, Interchangeability, Martingales.* Third Edition, Springer - Verlag, New York.

Chung K. L. (1953). Sur Les Lois De Probabilities Unimodales. *C. R. Acad. Sci. Paris* **236** 583–584.

Chung, K. L. (1976). *A course in Probability Theory.* Academic Press, New York.

Csorgo, S. Deheuvels, P. and Mason, D. M. (1985). Kernel estimates of the tail index of a distribution. *Ann. Statist.* **13** 1050–1077.

Damien, Paul and Laud, Purushottam W. and Smith, Adrian, F. M. (1995). Approximate random variate generation from infinitely divisible distributions with applications to Bayesian inference. *J. Roy. Statist. Soc. Ser.* B **57**, 3, 547–563.

Damien, Paul and Laud, Purushottam W. and Smith, Adrian, F. M. (1996). Implementation of Bayesian non-parametric inference based on beta processes. *Scan. J. Statist.*, **23**, 1, 27-36.

Daniel, Edward J. and Katti, S. K. (1987). On the Poisson characterization through infinite divisibility. *J. Indian Statist. Assoc.* **25**, 97–98.

Danielsson, J., de Haan, L., Peng, L. and de Vries, C. G. (2001). Using a bootstrap method to chhose the sampling fraction in tail index estimation. *Jour. Mult. Anal.*, **76**, 2, 226–248.

Datta, S. and McCormick, W. P. (1998). Inference for the tail parameters of a linear process with heavy tail innovations. *Ann. Inst. Statist. Math.* **50**, 337–359.

de Haan, L. and Resnick, S. I. (1980). A simple asymptotic estimate for the index of a stable distribution. *J. Roy. Statist. Soc. Ser.* B **42**, 83–87.

de Haan, L and Resnick, S. I. (1998). On asymptotic normality of the Hill estimator. *Comm. Statist. Stochastic Models*, **14**, 4, 849–866.

de Haan, L and Peng, L. (1998). Comparison of tail index estimates. *Statistica Neerlandica*, **92**. 1, 60–70.

de Haan, L. and Pereira, T. Themido (1999). Estimating the index of a stable distribution. *Statist. Probab. Lett.* **4**, 1, 39–55.

de Vries, C. G. (1991). On the relation between GARCH and stable processes. *J. Econometrics* **48**, 3, 313–324.

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

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.

Dekkers, A. L. M. and de Haan, L. (1993). Optimal choice of sample fraction in extreme value estimates. *Jour. Multi. Anal.* **2**, 173–195.

Dharmadhikari, Sudhakar and Joag - Dev, Kumar (1988). *Unimodality, Convexity and Applications.* Academic Press, Inc. Boston.

Drees, H., de Haan, L. and Resnick, S. (2000). How to make a Hill plot. *Ann. Statist.*, **28**, 1, 254–274.

DuMouchel, W. H. (1973). On the asymptotic normality of the maximum likelihood estimate when sampling from a stable distribution. *Ann. Statist.* **1**, 948–957.

DuMouchel, W. H. (1973). Stable distributions in statistical inference. I. Symmetric stable distributions compared to other symmetric long-tailed distributions. J. Amer. Statist. Assoc. **68**, 469–477.

DuMouchel, W. H. (1975). Stable distributions in statistical inference. II. Information from stably distributed samples. J. Amer. Statist. Assoc. **70**, 386–393.

DuMouchel, W. H. (1983). Estimating the stable index α in order to measure tail thickness: A Critique. *Ann. Statist.* **11**, 1019–1031.

Dwass, M. and Teicher, H. (1957). On infinitely divisible random vectors. *Ann. Math. Statist.* **28**, 461–470.

Embrechts, P., Goldie, C. M. and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility. *Z. Wahr. ver. Gebiete*, **49**, 3, 335–347.

Falk, M. (1995). On testing the extreme value index via the POT - method. *Ann. Statist.* **23**, 2013–2035.

Falk, M. and Marohn, F. (1997). Efficient estimation of the shape parameter in Pareto models with partially known scale. *Statist. Decision* **15** 229–239.

Fama, E. and Roll, R. (1971). Parameter estimates for symmetric stable distributions. *J.*

Amer. Stat. Assoc., **66**, 334, 331–338.

Feller, William (1996). *An Introduction to Probability Theory and its Applications.* Vol II. John Wiley & Sons, Inc. New York.

Feuerverger, Andrey and McDunnough, Philip (1981). On efficient inference in symmetric stable laws and processes. *Statistics and Related Topics* (Ottawa, Ont., 1980) pp.109–122, North Holland, Amsterdam New York.

Feuerverger, A. and McDunnough, P. (1981). On the efficiency of empirical characteristic function procedures. *J. Roy. Stat. Soc. Ser. B*, **43**, 1, 20–27.

Ferguson, T. and Phadia, E. G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.* **7**, 163–186.

Gamrowski, B. and Rachev, S. T. (1995). Stable models in testable asset pricing, in Approximation, Probability, and Related Fields. G. Aneastassiou and S. T. Rachev, eds.

Gnedenko, B. V. and Kolmogorov (1954). *Limit Distributions for Sums of Independent Random Variables.* Addison - Wesley, Cambridge.

Goldie, Charles (1967). A class of infinitely divisible random variables. *Proc. Cambridge Philos. Soc.* **63** 1141–1143.

Goldie, Charles M. and Kluppelberg, Claudia (1998). Subexponential distributions. *A practical guide to heavy tails.* 435–459, Birkhauser, Boston.

Goldie, C. M. and Smith, R. L. (1987). Slow variation with remainder: theory and applications. *Quarterly J. of Math.*, Oxford 2nd series, **38**, 45–71.

Goovaerts, M. J., D'Hooge, L., de Pril, N. (1977). On a class of generalized Γ-convolutions. I *Scand. Actuar. J.* **1**, 21–30.

Goovaerts, M. J., D'Hooge, L., de Pril, N. (1977). On the infinite divisibility of the product of two Γ-distributed stochastical variables. *Appl. Math. Comput.* **3**, 2, 127–135.

Goovaerts, M. J., D'Hooge, L., de Pril, N. (1978). On the infinite divisibility of the ratio of two gamma-distributed variables. *Stoch. Proc. Appl.*. **7**, 3, 291–297.

Griffiths, R. C.(1970). Infinitely divisible multivariate gamma distributions. *Sankhyā Ser. A* **32**, 393–404.

Griffiths, R. C. (1984). Characterization of infinitely divisible multivariate gamma distributions. *J. Multivariate Anal.* **15**, 1, 13–20.

Griffiths, R. C. and Milne, R. K. (1987) A class Of infinitely divisible multivariate negative binomial distributions. *J. Multivariate Anal.* **22**, 1, 13–23.

Grosswald, E. (1976). The Student t-distribution for odd degrees of freedom is infinitely divisible. *Ann. Probab.* **4**, 680–683.

Grosswald, E. (1976). The Student t-distribution of any degree of freedom is infinitely divisible. *Z. Wahrsch. Verw. Gebiete* **36**, 2, 103–109.

Hall, P. (1981). Order of magnitude of moments of sums of random variables. *J. London Math. Soc.* (2) 24, 3, 562–568.

Hall, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc.* **B44**, 37–42.

Hall, P. (1990). Using the bootstrap to estimate mean square error and select smoothing parameter in nonparametric problem. *J. Mult. Anal.* 32, 177–203.

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

Hansen, B. G. (1990). *Monotonicity Properties of Infinitely Divisible Distributions.* CWI Tract, 69. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam.

Hartman, P. and Wintner, A. (1992). On the infinitesimal generators of integral convolutions. *Amer. J. Math.*, **64**, 273–398.

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

Höpfner, R. (1997). On tail parameter estimation in certain point process models. *J. Statist. Planning and Inf.* **60**, 169–187.

Höpfner, R. (1997) Two comments in parameter estimation in stable processes. *Math. Meth. Statist.* **6**, 125–134.

Höpfner, R. (1998). Estimating a stability parameter: asymptotics and simulations. *Statistics* **30**, 291–305.

Höpfner, R. and Jacod, J. (1994). Some remarks on the joint estimation of the index and the scale parameter for stable processes. *Asymptotic Statistics* (Prague, 1993), 273–284, Contrib. Statist., Physics, Heidelberg, 1994.

Höpfner, R. and Rüschendorf, L. (1996). Comparison of estimators in stable models. *Math. Comput. Modelling* **29**, 10–12, 145–160.

Höpfner, R. (1997). Two comments on parameter estimation in stable processes. *Math. Methods Statist.* **6**, 1, 125–134.

Hjort, N. L. (1990). Nonparametric Bayes estimates based on beta processes in models for life history data. *Ann. Statist.* **18**, 1254–1294.

Horn, R. A. and Steutel, F. W. (1977/78). On multivariate infinitely divisible distributions. *Stochastic Process. Appl.* **6**, 2, 139–151.

Hougaard, Philip (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**, 2, 387–396.

Hougaard, Philip (1988). Correction to: "Survival models for heterogeneous populations".

Biometrika **75**, 2, 395.

Hsing, Tailen. (1991). On tail index estimation using dependent data. *Ann. Statist.* **19**, 1547–1569.

Hsing, Tailen (1995). Limit theorems For stable processes with application to spectral density estimation. *Stochastic Process. Appl.* **57**, 1, 39–71.

Ibragimov, I. A. (1956). On the composition of unimodal distributions. *Theor. Probab. Appl.* **2**, 117–119.

Ibragimov, I. A. (1957). A remark on probability distributions of class *L*. *Theor. Probab. Appl.* **2**, 117–119.

Ibragimov, I. A. Černin, K. E. (1959). On the unimodality of stable laws. (in Russian) *Teor. Ver.i Primenen.* **4**, 453–456.

Jakeman, E. and Pusey, P. N. (1976). A model for non-Rayleigh sea echoes, *IEEE Trans. on Antennas and Propagation.* 24, 6, 806–814.

Janicki, Aleksander and Weron, Aleksander (1994). Can one see α-stable variables and processes? *Statist. Sci,* **9**, 1, 109–126.

Joe, Harry (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statist. Probab. Lett.* **9**, 1, 75–81.

Joe, H. (1993). Parametric families of min-stable distributions with given marginals. *J. Multivariate Anal.*, **46**, 262–282.

Joe, Harry and Hu, Taizhong (1996). Multivariate distributions from mixtures of max-infinitely divisible distributions. *J. Multivariate Anal.* **57**, 2, 240–265.

Karatzas, Ioannis, Rajput, Balram, S. and Taqqu, Murad, S. (1998). (Eds). *Stochastic Process and Related Topics.* Birkhäuser, Boston.

Katti, S. K. (1967). Infinite divisibility of integer-valued random variables. *Ann. Math. Statist.* **38**, 1306–1308.

Katti, S. K. (1977). Infinite divisibility of discrete distributions. III. *Analytic Function Methods in Probability Theory. Colloq. Math. Soc. Janos Bolyai* **21**, 165–171.

Keilson, J. and Steutel, F. W. (1972). Families of infinitely divisible distributions closed under mixing and convolution. *Ann. Math. Statist.* **43**, 242–250.

Klüppelberg, Claudia and Mikosch, Thomas (1994). Some limit theory for the self-normalized periodogram of stable processes. *Scand. J. Statist.* **21**, 4, 485–491.

Klüppelberg, Claudia and Mikosch, Thomas (1993). Spectral estimates and stable Processes. *Stochastic Process. Appl.* **47**, 2, 323–344.

Klüppelberg, Claudia and Mikosch, Thomas.(1996). The integrated periodogram for stable processes. *Ann. Statist.* **24**, 5, 1855–1879.

Knight, K. (1989). Limit theory for autoregressive parameter estimates in an infinite-variance random walk. *Can. J. Statist.* **17**, 261–278.

Kogan, S. M. and Manolakis, D. G. (1996). Signal modeling with self-similar α-stable processes: the fractional Levy stable motion model. *IEEE Trans. on Signal Processing*, **44**, 4, 1006–1010.

Kogan, S. M. and Williams, D. B. (1998). Characteristic function based estimation of stable distribution parameters. In *A practical guide to heavy tails*, (Eds. Adler, R. J., Feldman, R. E. and Taqqu, M. S.) pp. 311–338. Birkhauser, New York.

Kolmogorov, A. N. (1956). Two uniform limit theorems for sums of independent summands. *Teor. ver. Primenen.*, 1, 426–436.

Konev, V. V. and Pergamenshchikov, S. M. (1996). On asymptotic minimaxity of fixed accuracy estimators for autoregression parameters. I. Stable Process. *Math. Methods Statist.* **5**, 2, 125–153.

Koutrouvelis, I. A. (1980). Regression-type estimation of the parameters of stable laws. *J. Amer. Stat. Assoc.*, **75**, 372, 919–928.

Koutrouvelis, I. A. (1981). An iterative procedure for the estimation of the parameters of stable laws. *Comm. in Statist.-Simulation and Computation*, 10, 1, 17–28.

Koutrouvelis, Ioannis A. and Bauer, David F. (1982). Asymptotic distribution of regression-type estimators of parameters of stable laws. *Comm. Statist. A-Theory Methods* **11**, 23, 2715–2730.

Laha, R. G. (1961). On a class of unimodal distributions. *Proc. Amer. Math. Soc.* **12**, 181–184.

LePage, Raoul (1980). Multidimensional infinitely divisible variables and processes. II. *Probability in Banach Spaces*, III. (Medford, Mass.,) 279–284, *Lecture Notes in Math.*, 860, Springer, New York.

LePage, Raoul and Podgórski, Krzysztof (1996). Resampling permutations in regression without second moments. J. *Multivariate Anal.* **57**, 1, 119–141.

Lévy, P. (1937). *Théorie de ĺ Addition des Variables Aléatoires* Gauthier, Villars, Paris.

Linde, Werner (1986). *Probability in Banach Spaces-Stable and Infinitely Divisible Distributions.* Second Edition. John Wiley, Chichester.

Liu, S. M. and Brorsen, B. W. (1995). Maximum likelihood estimation in GARCH-stable model. *Jour. Appl. Econometrics* 10, 273–285.

Loéve, M. (1955). *Probability Theory.* Van Nostrand, New York.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *J. Business*, **36**, 394–419.

Marohn, F. (1999). Estimating the index of a stable law via the Pot-method. *Stat. Prob.*

Lett. **41**, 413–423.

Marshall, A. W. and Olkin (1988). Families of multivariate distributions. *J. Amer. Stat. Assoc.* **83**, 834–841.

Mason, D. M. (1982). Laws of large numbers for sums of extreme values. *Ann. Probab.*, **10**, 3, 754–764.

Matthes, Klaus; Kerstan, Johannes and Mecke, Joseph. (1978). *Infinitely Divisible Point Processes.* John Wiley & Sons, New York.

McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Comm. in Stat.-Simulation and Computation*, **15**, 4, 1109–1138.

McCulloch, J. H. (1996a). On the parameterization of the afocal stable distributions. *Bull. of London Math. Society*, **28**, 651–655.

McCulloch, J. H. (1996b). Financial applications of stable distributions. In G. S. Maddala and C. R. Rao, eds., *Statistical Methods in Finance* Handbook of Statistics, 14, 393–425. Elsevier Science, Amsterdam.

McCulloch, J. H. (1997). Measuring tail thickness in order to estimate the stable index. A critique. *J. Business Econom. Statist.*, **15**, 1, 74–81.

McCulloch, J. H. (1998a). Linear regression with stable disturbances. In *A practical guide to heavy tails* (Eds. Adler, R. J., Feldman, R. E. and Taqqu, M. S.) pp 359–378. Birkhauser, New York.

McCulloch, J. H. (1998b). Numerical approximation of the symmetric stable distribution and density. In *A practical guide to heavy tails* (Eds. Adler, R.J., Feldman, R. E. and Taqqu, M. S.). pp. 489–499. Birkhauser, New York.

McCulloch, J. W. and Panton, D. B. (1998). Tables of the maximally-skewed stable distributions and densities. In *A practical guide to heavy tails* (Eds. Adler, R.J., Feldman, R. E. and Taqqu, M. S.). pp. 501–508. Birkhauser, New York.

McKenzie, E. (1982). Product autoregression: a time-series characterization of the gamma distribution.*J. Appl. Prob.*, **19**, 463–468.

Mesalkin, L. D. (1960). Approximation of a multinomial distribution by infinitely divisible laws. *Teor. Ver. Primenen.*, **5**, 114–124.

Mijnheer, J. (1997). Asymptotic inference for AR(1) processes with (nonnormal) stable errors. Proceedings of the XVII Seminar on Stability Problems for Stochastic Models, Part II (Kazań, 1995). *J. Math. Sci.* **83**, 3, 401–406.

Mijnheer, J. (1997). Asymptotic inference for AR(1) processes with (nonnormal) stable errors.III. Heavy Tails and Highly Volatile Phenomena. *Comm. Statist. Stochastic Models* **13**, 4, 661–672.

Mijnheer, J. (1998). Asymptotic inference for Ar(1) processes with (nonnormal) stable errors. IV. A note on the case of a negative root. Proceedings of the 18th Seminar on

Stability Problems for Stochastic Models, Part II (Hajdūszobozló, 1997). *J. Math. Sci.* **92**, 4, 4035–4037.

Miller, G. (1978). Properties of certain symmetric stable distributions. *J. Mult. Anal.* **8**, 3, 346–360.

Mittnik, S. and Paolella, M. S. (1999). A simple estimator for the characteristic exponent of the stable Paretian distribution. *Math. Comput. Modelling* **29**, 10–12, 161–176.

Mittnik, S. and Rachev, S. T. (1993). Modeling asset returns with alternative stable distributions. *Econometric Rev.* **12**, 3, 261–389.

Moothathu, T. S. K. (1985). On a characterization property of multivariate symmetric stable distributions. *J. Indian Statist. Assoc.* **23**, 83–88.

Nikias, C. L. and Shao, M. (1995). *Signal Processing with Alpha-Stable Distributions and Applications.* John Wiley, New York.

Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions, *Stochastic Models*, **13**, 759–774.

Nolan, J. P. (1998). Parametrizations and modes of stable distributions. *Stat. Prob. Letters*, **38**, 2, 187–195.

Nolan, J. P. (1998). Univariate stable distributions: parameterizations and software. In *A practical guide to heavy tails* (Eds. Adler, R. J., Feldman, R. E. and Taqqu, M. S.) pp 359–378. Birkhauser, New York.

Nolan, J. P. (2001). Maximum likelihood estimation and diagnostics for stable distribution. In *Levy Processes* (Eds, Barndorff-Nielsen, O., Mikosch, T, and Resnick, S. I. pp 379–400. Birkhauser, New York.

Nolan, J. P., Panorska, A. K. and McCulloch, J. H. (1996). Estimation of stable spectral measures. Tech. Report. Ohio State University.

Officer, R. R. (1972). The distribution of stock returns. *Jour. Amer. Stat. Assoc.* 67, 807–812.

Oh, C. S. (1994). *Estimation of time varying term premia of U.S. Treasury Securities: using a STARCH model with stable distributions.* Ph.D. dissertation, Ohio State University.

Paulson, A. S. and Delehanty, T. A. (1985). Modified weighted squared error estimation procedures with special emphasis on the stable laws. *Comm. in Statistics-Simulation and Computation*, **14**, 4, 927–972.

Paulson, A. S., Holcomb, E. W. and Leitch, R. A. (1975). The estimation of the parameters of the stable laws. *Biometrika*, **62**, 163–170.

Pickands, III., J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, **3**, 1, 119–131.

Pierre, P. A. (1971). Infinite divisible distributions, conditions for independence, and central limit theorems. *J. Math. Anal. Appl.* **33**, 341–354.

Pitman, J. and Yor, Marc (1981). Bessel processes and infinitely divisible laws. Stochastic Integrals (*Proc. Sympos. Univ. Durham, Durham,* 1980), 285–370, *Lecture Notes in Math.*, 851, Springer, Berlin, 1981.

Pointer, S. (1996). Evidence of non-Gaussian scaling behavior in heterogeneous sedimentary formations. *Water Resources Research,* **32**, 5, 1183–1195.

Press, S. J. (1972). Estimation of univariate and multivariate stable distributions. *J. Amer. Stat. Assoc.*, **67**, 340, 842–846.

Prohorov, Yu. V. (1960). On a uniform limit theorem of A. N. Kolmogorov, *Teor. ver. Primenen..*, 103–113.

Rachev, S. T. and Resnick, S. I. (1991). Max-geometric infinite divisibility and stability. *Comm. Statist. Stochastic Models* **7**, 191–218.

Resnick, S. and Rootzen, H. (2000). Self-similar communication models and very heavy tails. *Ann. appl. Probab.*, **10**, 3, 753–778.

Resnick, S. and Stă rică, C. (1995). Consistency of Hill's estimator for dependent data. *Jour. Appl Probab.*, **32**, 1, 139–167.

Resnick, S. and Stă rică, C. (1997a). Asymptotic behavior of Hill's estimator for autoregression data. heavy tails and highly volatile phenomena. *Comm. Stat. Stoch. Models*, **13**, 4, 703–721.

Resnick, S. and Stă rică, C. (1997b). Smoothing the Hill estimate. *Adv. App. Probab.*, **29**, 1, 271–293.

Resnick, S. and Stă rică, C. (1998). Tail index estimation for dependent data. *Ann. App. Probab.*, **8**, 4, 1156–1183.

Resnick, S. and Stă rică, C. (1999). Smoothing the moment estimator of the extreme value parameter. *Extremes* **1**, 3, 263–293.

Rohatgi, Vijay K., Steutel, F. W. and Székely, Gábor, J. (1990). Infinite Divisibility of Products and Quotients of i.i.d. Random Variables. *Math. Sci.* **15**, 1, 53–59.

Rootzen, H., Leadbetter, M. R. and de Haan, L. (1990). Tail and quantile estimation for strongly mixing stationary sequences. *Preprint* Center for Stochastic Process. Univ. of North Carolina.

Rosínski, J. (1991). On a class of infinitely divisible processes represented as mixtures of Gaussian processes. *Stable Processes and Related Topics* (Ithaca, NY, 1990), 27–41, *Progr. Probab.*, 25, Birkhäuser, Boston.

Rvaĉeva, E. L. (1954). On domains of attraction of multidimensional distributions. *Selected Translations Math. Statists. Probab.*, Vol 2, 183–205. American Mathematical Society, Providence, Rhode Island.

Samorodnitsky, G. and Taqqu, M. (1993). Stochastic monotonicity and Slepian-type inequalities for infinitely divisible and stable random vectors. *Ann. Prob.*, 21, 1, 143–160.

Samorodnitsky, Gennady and Taqqu, Murad, S. (1994). *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York.

Sato, K. (1994). Multimodal convolutions of unimodal infinitely divisible distributions. *Theory Probab. Appl.* **39**, 2, 336–347.

Sato, K. (1999). *Levy processes and infinitely divisible distributions.* Cambridge University Press.

Sclove, Stanley L. (1981). Some recent statistical results for infinitely divisible distributions. In *Statistical Distributions in Scientific Work.* Vol. 4 (Trieste, 1980), 267–280 NATO *Adv. Study Fust. Ser. C. Math. Phys. Sci.*, 79, Reidel, Dordrecht, Boston.

Sclove, Stanley L. (1983). Some aspects of inference for multivariate infinitely divisible distribution. *Statist. Decisions* **1**, 3, 305–321.

Sharpe, M. J. (1969). Zeroes of infinitely divisible densities. *Ann. Math. Statist.* **40**, 1503–1505.

Steutel, F. W. (1967). Note on the infinite divisibility of exponential mixtures. *Ann. Math. Statist.* **38** 1303–1305.

Steutel, F. W.(1968). A class of infinitely divisible mixtures. *Ann. Math. Statist.* **39**, 1153–1157.

Steutel, F. W. (1969). Note on completely monotone densities. *Ann. Math. Statist.* **40** 1130–1131.

Steutel, F. W. (1970). *Preservation of Infinite Divisibility Under Mixing and Related Topics.* Mathematical Centre Tracts, 33. Mathematisch Centrum, Amsterdam.

Steutel, F. W. (1973). Some recent Results in infinite divisibility. *Stoch. Processes Appl.* **1**, 125–143.

Steutel, F. W. (1973/74). On the tails of infinitely divisible distributions. *Z. Wahr. Verw. Gebiete* **28**, 273–276.

Steutel, F. W. (1979). Infinite divisibility in theory and practice. *Scan. J. Statist.* **6**, 2, 57-64.

Steutel, F. W. (1979). Infinite divisibility of mixtures of gamma distributions. *Colloq. Math. Soc. János Bolyai* **21**, 345–357.

Stoyanov, Jordan, M. (1987). *Counter Examples in Probability.* John Wiley & Sons, Ltd, Chichester.

Stoyanov, J., Miraychiwiski, I., Ignatov, Z. and Tanushev, M. (1989). *Exercise Manual in Probability Theory.* Kluwer Academic Publishings Group, Dordrecht.

Stuck, B. W. and Kleiner, B. (1974). A statistical analysis of telephonic noise. *Bell Systems*

Technical Journal, **53**, 1262–1320.

Takano, Katsuo (1994). On infinitely divisible distributions of modified Bessel functions. *Bull. Fac. Sci. Ibaraki Univ. Ser.* **A 26**, 25–28.

Talwalker, S. (1979). A note on the characterization of the multivariate normal distribution. *Metrika*, **26**, 1, 25–30.

Thorin, O. (1977). On the infinite divisibility of the lognormal distribution. *Scand. Actuar. J.* **3**, 121–148.

Thorin, Olaf (1977). On the Infinite Divisibility of the Pareto distribution. *Scand. Actuar. J.* **1**, 31–40.

Thorin, Olaf (1978). Proof of a conjecture of L. Bondesson concerning infinite divisibility of powers of a gamma variable. *Scand. Actuar. J.* **3**, 151–164.

Thorin, Olaf (1978). An extension of the notion of a generalized Γ convolution. *Scand. Actuar. J.* **3**, 141–149.

Thornton, J. C. and Paulson, A. S. (1977). Asymptotic distribution of characteristic function–based estimators for the stable laws. *Sankhyā Ser. A*, **39**, 4, 341–354.

Tikhov, M. S. (1992). Asymptotics of T-Estimators. *Theory Probab. Appl.* **37**, 4, 644–657.

Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and Stability. Stable laws and their application.* Modern probability and Statistics, VSP, Utrecht.

van Harn, Klaas (1978). *Classifying Infinitely Divisible Distributions by Functional Equations.* Mathematical Centre Tracts, 103. Mathematisch Centrum, Amsterdam.

Veeh, J. (1982). Infinitely divisible measures with independent marginals. *Z. Wahrsch. Verw. Gebiete* **61**, 3, 303–308.

Walker, S. and Damien, Paul. (2000). Representations of Levy processes without Gaussian components. *Biometrika*, **87**, 2, 477–483.

Wang, Y. H. (1976). A functional equation and its application to the characterization of the Weibull and stable distributions. *J. Appl. Probability*, **13**, 2, 385–391.

Warde, W. D. and Katti, S. K. (1971). Infinite divisibility of discrete distributions. II. *Ann. Math. Statist.* **42**, 1088–1090.

Weron, Aleksander and Weron, Rafal (1995). Computer simulation of Lévy α-Stable Variables and Processes. Chaos–The Interplay Between Stochastic and Deterministic Behavior (Karpacz, 1995), 379–392, *Lecture Notes in Phys.*, 457, Springer, Berlin, 1995.

Wesolowski, Jacek (1993). Multivariate infinitely divisible distributions with the Gaussian second order conditional structure. In *Stability Problems for Stochastic Models* (Suzdal, 1991), 180–183, *Lecture Notes in Math.*, 1546, Springer, Berlin, 1993.

Willekens, E. (1987). On the supremum of an infinitely divisible process. *Stoch. Proc. Appl.*, **26**, 1, 173–175.

Wintner, A. (1956). Indefinitely divisible symmetric laws and normal stratifications. *Pub. Stat. Univ. de Paris* **6**, 327–336.

Wolfe, S. J. (1971). On moments of infinitely divisible distributions. *Ann. Math. Statist.*, **42**, 2036–2043.

Wolfe, S. J. (1981). On the Unimodality of infinitely divisible distribution functions. II. *Analytic Methods in Probability Theory* (Oberwolfach, 1980) 178–183, *Lecture Notes n Math.*, 861, Springer, New York.

Worsdale, Graham, J. (1975). Tables of cumulative distribution functions for asymmetric stable distributions. *J. Roy. Statist. Soc. Ser. C. Appl. Statist.* **24**, 123–131.

Yamazato, M. (1975). Some results on infinite divisible distributions of class L with applications to branching processes. *Sci. Rep. Tokyo Kyoiku Daigaku Sect.* A, **13**, 347–365, 133–139.

Yamazato, M. (1982). On strongly unimodal infinitely divisible distributions. *Ann. Probab.*, **10**, 3, 589–601.

Yamazato, M. (1996). On the strong unimodality of infinitely divisible distributions of class CME. *Theory Probab. Appl.* **40**, 3, 518–532.

Yanushkyavichene, O. L. (1981). Investigation of some estimates of the parameters of stable distributions. *Litovsk. Math. Sb.* **21**, 4, 195–209.

Zarepour, M. and Knight, K. (1999). Bootstrapping unstable first order autoregressive process with errors in the domain of attraction of stable law. *Comm. Statist. Stochastic Models* **15**, 1, 11–27.

Zeng, Wei Bin (1995). On characterization of multivariate stable distributions via random linear statistics. *J. Theoret. Probab.* **8**, 1995, 1, 1–15.

Zolotarev, V. M. (1981). Integral transformations of distributions and estimates of parameters of multidimensional spherically symmetric stable laws. *Contributions to Probability*, 283–305, Academic Press, New York-London, 1981.

Zolotarev, V. M. (1986). *One-Dimensional Stable Distributions.* Vol 65 of "Translations of Mathematical Monographs, American Mathematical Society. Translation from the Original 1983 Russian Edition.