INFINITELY DIVISIBLE DISTRIBUTIONS
Introduction

In these notes we discuss basic properties of infinitely divisible and stable distributions on $\mathbb{R}$. Infinitely divisible stochastic processes and domains of attraction are some important topics which will not be discussed here. They will be discussed in a future volume. Also, infinitely divisible and stable distributions on a Banach space will be studied in Volume 2 of these notes.

Measure theoretic preliminaries: if μ and ν are Borel probability measures on $\mathbb{R}$ we define the convolution $\mu * \nu$ by $(\mu * \nu)(E) = \int\limits_{\mathbb{R}} \mu(E-x)d\nu(x)$ where $E-x$ stands for $\{y-x : y \in E\}$. Noting that $I_{\{(x,y)\in\mathbb{R}^2 : x+y\in E\}}$ is a Borel measurable function on $\mathbb{R}^2$ for every Borel set E in $\mathbb{R}$ we conclude that $\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} I_{\{(x,y)\in\mathbb{R}^2:x+y\in E\}} d\mu(x)d\nu(y) = \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} I_{\{(x,y)\in\mathbb{R}^2:x+y\in E\}} d\nu(y)d\mu(x)$ which shows that $\mu*\nu = \nu*\mu$. Of course, $\mu*\nu$ is also a probability measure.

Notations: $\xrightarrow{w}$ denotes weak convergence (or convergence in distribution), $\xrightarrow{p}$ denotes convergence in probability and $X \overset{d}{=} Y$ means X and Y have the same distribution.

Suppose $\mu_1 = \mu$ and $\mu_{n+1} = \mu_n * \mu, n = 1, 2, \ldots$. We call μ_n the $n-$ fold convolution of μ with itself. If $\mu_n = \lambda$ then μ is called an $n-th$ root of λ. For a given probability measure λ and a given integer n there may be no $n-$ th root. [Examples will be given later]

A useful tool in studying roots of probability measures is the characteristic function. We shall write $\hat{\mu}$ for the characteristic function of μ. Recall that $\hat{\mu}(t) = \int\limits_{\mathbb{R}} e^{itx} d\mu(x)$. We assume that the reader is familiar with basic properties of characteristic functions. We reformulate the concept of $n-$ th root as follows:

a characteristic function ϕ has an $n-$ th root if there exists another characteristic function ψ such that $\psi^n(t) = \phi(t)\ \forall t \in \mathbb{R}$.characteristic functions then $\psi_1 \equiv \psi_2$: indeed

Uniqueness: if $\psi_1^n(t) = \psi_2^n(t)$ where ψ_1 and ψ_2 are both

characteristic functions and ψ_1 never vanishes then $\frac{\psi_1(t)}{\psi_2(t)}$ is an $n-$ th root of unity for each t. By continuity it follows that $\frac{\psi_1(t)}{\psi_2(t)}$ is an $n-$ th root of unity which is independent of t. However $\frac{\psi_1(0)}{\psi_2(0)} = 1$ so $\frac{\psi_1(t)}{\psi_2(t)} \equiv 1$.

A probability measure μ (or its characteristic function ϕ) is called *infinitely divisible* (i.d.) if it has an $n-$ th root for every positive integer n. We call a random variable infinitely divisible if the induced measure is.

We can also formulate these concepts in terms of random variables: X is i.d. if, for each n, there exist independent identically distributed (i.i.d.) random variable $X_{1,n}, X_{2,n}\ ,\ldots, X_{n,n}$ (not necessarily on the same probability space) such that $X_{1,n} + X_{2,n} + \ldots + X_{n,n} \stackrel{d}{=} X$.

Examples:

1. Let $\phi(t) = e^{ict}e^{-t^2\sigma^2/2}$ where $c \in \mathbb{R}$ and $\sigma > 0$. ϕ is the characteristic function of normal distribution with mean c and variance σ^2. If $\phi_n(t) = e^{ict/n}e^{-t^2\sigma^2/(2n)}$ then ϕ_n is the characteristic function of normal distribution with mean c/n and variance σ^2/n. Also $\phi_n^n = \phi$. Hence ϕ is i.d..

2. Uniform distribution on an interval is not i.d.. We shall show that *no non-constant bounded random variable can be i.d.*! suppose $X_1, X_2, \ldots, X_n$ are i.i.d. and $X_1 + X_2 + \ldots + X_n \stackrel{d}{=} X$ where X is a bounded random variable, say with $|X| \leq M$ a.s.. Then $Var(X_1) = \frac{1}{n}Var(X)$. We claim that $|X_1| \leq \frac{M}{n}$ almost surely (a.s.). In fact, $0 = P\{X > M\} \geq P\{X_1 > \frac{M}{n}, X_2 > \frac{M}{n}, \ldots, X_n > \frac{M}{n}\} = P^n\{X_1 > \frac{M}{n}\}$ so $P\{X_1 > \frac{M}{n}\} = 0$; similarly, $P\{X_1 < -\frac{M}{n}\} = 0$ so $|X_1| \leq \frac{M}{n}$ a.s.. This implies that $Var(X_1) \leq \frac{M^2}{n^2}$ and hence $Var(X) \leq \frac{M^2}{n}$. If X is i.d. then this inequality holds for every n so $Var(X) = 0$ which is a contradiction.

3. Let X be Cauchy random variable. The characteristic function of X is given by $Ee^{itX} = e^{-|t|}$. It follows that $E^{e^{it(\frac{X}{n})}} = Ee^{i(\frac{t}{n})X} = e^{-\frac{|t|}{n}}$. Hence $e^{-\frac{|t|}{n}}$ is a characteristic function whose $n-$ th power is the characteristic function of X.

4. Let X take the values 0 and 1 with probability $p$and $1-p$. Suppose there exist i.i.d. random variables Y and Z such that $X \stackrel{d}{=} Y + Z$. Then $1 = P\{Y + Z \in \{0,1\})\} = \displaystyle\int P\{Z \in \{0,1\} - y\}d\mu(y)$ where μ is the measure induced by Y. Since the integrand takes values on $[0,1]$ it follows that $P\{Z \in \{0,1\} - y\} = 1$ almost everywhere (a.e.) with respect to μ. In particular there exists a real number y such that $P\{Z \in \{-y, 1-y\}\} = 1$. Hence, we also have$P\{Y \in \{-y, 1-y\}\} = 1$. Note that $P\{Y = -y\}$ and $P(Y = 1-y\}$ must both be non-zero. (Otherwise Y, hence Z, would be constants and so would be X, a contradiction). If $u \in \{-y, 1-y\} + \{-y, 1-y\}$ then $P\{X = u\} = P\{Y + Z = u\} > 0$. [If $u = u_1 + u_2$ where $u_1, u_2 \in \{-y, 1-y\}$ then $P\{X = u\} \geq P\{Y = u_1, Z = u_2\} \geq P\{Y = u_1\}P\{Z = u_2\} > 0$]. Hence the possible values of X are the points of $\{-y, 1-y\} + \{-y, 1-y\}$ which implies $\{-y, 1-y\} + \{-y, 1-y\} = \{0,1\}$. That there is no such real number y is obvious.

5. Discrete random variables may be i.d.. Let X have the Poisson distribution with parameter λ. Then $P\{X = n\} = e^{-\lambda}\frac{\lambda^n}{n!}, n = 0, 1, 2, \ldots$ and

$Ee^{itx} = \sum_{n=0}^{\infty} e^{itn} e^{-\lambda} \frac{\lambda^n}{n!} = e^{\lambda(e^{it}-1)}$. Hence Poisson distribution with parameter λ/n is an $n-$ th root of the one with parameter λ.

6. Geometric distribution is i.d.: if $P\{X = n\} = (1-p)p^n, n = 0, 1, 2, \ldots$ where $0 < p < 1$ then $Ee^{itx} = \frac{1-p}{1-pe^{it}}$. We leave it as an exercise to show that if $P\{X_n = k\} = \binom{n+k-1}{k} p^k (1-p)^n$ then $Ee^{itx} = (Ee^{itX_n})^n$. More generally negative Binomial distribution is i.d. by a similar argument.

7. Let X have the Gamma distribution with parameters a and b. Then X has density $f(x) = \frac{b^a e^{-bx} x^{a-1}}{\Gamma(a)}$ where $a, b > 0$. Then $Ee^{itx} = \frac{1}{(1-iat)^b}$ where $(1-iat)^b$ is defined as $e^{bLog(1-iat)}$, Log being the principle branch of logarithm. It is now obvious that X is i.d.

8. Suppose $f(x) = (1-|x|)^+$. f is the density function of a random variable X. We have $Ee^{itX} = 2\int_0^1 (1-x)\cos(tx)dx = 2\frac{1-\cos t}{t^2}$ for $t \neq 0$. It is shown below that an i.d. characteristic function has no zeros. Since $1-\cos t = 0$ for $t = 2\pi$ it follows that X is not i.d.. Also $2\frac{1-\cos t}{t^2}$ is a density function whose characteristic function is $f(x)$ which has zeros, hence $2\frac{1-\cos t}{t^2}$ is not an i.d. density function; the inversion theorem from Fourier analysis makes these points obvious].

It is now clear that it may be hard to determine if a given distribution is i.d.. For more information on infinite divisibility of specific distributions we refer the reader to the book "Infinite Divisibility Of Probability Distributions On The Real Line" by Steutel and van Harn.

Remark: constant random variables are i.d. and linear combinations of independent i.d. distributions are i.d. In particular, if μ and ν are i.d. so is $\mu * \nu$. (Equivalently, product of two i.d. characteristic functions if i.d.). Proofs of these are left to the reader. Also note that if X is i.d. with characteristic function ϕ and $\{X, Y\}$ is i.i.d. then the characteristic function of $X - Y$ is $|\phi|^2$. Hence $|\phi|^2$ is an i.d. characteristic function whenever ϕ is.

We now describe a procedure for generating i.d. distributions from arbitrary distributions.

Theorem 1

Let ϕ be any characteristic function and $\lambda > 0$. Then $e^{-\lambda(1-\phi(t))}$ is an i.d. characteristic function.

Proof: we have $e^{-\lambda(1-\phi(t))} = \lim_{n\to\infty}(1-\frac{\lambda(1-\phi(t))}{n})^n$ for each t. Note that $1-\frac{\lambda(1-\phi(t))}{n} = (1-\frac{\lambda}{n})+\frac{\lambda}{n}\phi(t)$. Hence, if $\phi(t) = \int e^{itx}d\mu(x)$ then $1-\frac{\lambda(1-\phi(t))}{n} = \int e^{itx}d\nu(x)$ where $\nu = (1-\frac{\lambda}{n})\delta_0+\frac{\lambda}{n}\mu$. Since ν is a probability measure it follows that $1-\frac{\lambda(1-\phi(t))}{n}$ is a characteristic function and so is $(1-\frac{\lambda(1-\phi(t))}{n})^n$. By the Continuity Theorem for characteristic functions it now follows that $e^{-\lambda(1-\phi(t))}$ is a characteristic function. This characteristic function is the $n-$ th power of $e^{-\frac{\lambda}{n}(1-\phi(t))}$, which is also a characteristic function. This completes the proof.

We may ask if the characteristic functions constructed above exhaust all i.d. ones. The answer is no. Let us show that the normal characteristic function $\phi(x) = e^{-x^2/2}$ is not of above type. If $e^{-x^2/2} = e^{-\lambda(1-\phi(x))}$ $\forall x$ then $x^2/2 = \lambda(1-\operatorname{Re}\phi(x))$. This is clearly a contradiction since the right side is bounded.

Exercise

Show that for each positive integer n there exists a characteristic function ϕ such that ϕ is not i.d. but it is the $k-$ the power of a characteristic function for each $k \leq n$.

Solution: $\phi = \psi_0^{n!}$ where ψ is a characteristic function which has a zero.

One of our main jobs is to characterize i.d. characteristic functions. This is done below in Levy - Khinchine Theorem. See also Schoenberg's theorem below. Before doing this we prove some basic facts about i.d. distributions.

We first recall an elementary result from complex analysis.

Lemma 2

Let $f : [a, b] \to \mathbb{C}$ be a continuous function such that $f(x) \neq 0$ for any $x \in [a, b]$. Then there exists a continuous function $g : [a, b] \to \mathbb{C}$ such that $f = e^g$. If $c \in [a, b]$ and $f(c) = e^z$ then we can choose g such that $g(c) = z$. With this condition g is unique. The conclusion also holds if the domain of f is $\mathbb{R}$.

Proof: there is no loss of generality in assuming that $f(a) = 1$. If $\rho = \inf\{|f(t)| : a \leq t \leq b\}$ then $0 < \rho \leq 1$. For $|z-1| \leq \frac{1}{2}$ let $l(z) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k}(z-1)^k$. Then l is analytic in $B(1, \frac{1}{2})$, continuous on the closure of this ball and $l'(z) = \sum_{k=1}^{\infty}(-1)^k(z-1)^{k-1} = \frac{1}{1-(1-z)} = \frac{1}{z}$. It follows that $(e^{-l(z)}z)' = 0$ in $B(1, \frac{1}{2})$ so $e^{-l(z)}z$ is a constant. Since $l(1) = 0$ we get $e^{-l(z)}z = 1$ or $e^{l(z)} = z$ $\forall z \in B(1, \frac{1}{2})$. Since f is uniformly continuous on $[a, b]$ there exists $\delta > 0$ such that $|f(t_1) - f(t_2)| < \frac{\rho}{2}$ if $|t_1 - t_2| \leq \delta$. Let $\{t_j : 0 \leq j \leq N\}$ be a partition of $[a, b]$ such that $t_{j+1} - t_j \leq \delta$ for $0 \leq j < N$. Define g as follows: $g(t) = l(f(t))$ if $t \in [t_0, t_1]$ and $g(t) = g(t_k) + l(\frac{f(t)}{f(t_k)})$ if $t_k \leq t \leq t_{k+1}$ for $k = 1, 2, \ldots, N-1$. Note that if $t_k \leq t \leq t_{k+1}$ then $\left|1 - \frac{f(t)}{f(t_k)}\right| = \frac{|f(t) - f(t_k)|}{|f(t_k)|} \leq \frac{|f(t) - f(t_k)|}{\rho} < \frac{1}{2}$ so $g(t)$ is well defined. Also note that g is continuous on $[a, b]$ and $e^{g(t)} = e^{g(t_k)} \frac{f(t)}{f(t_k)}$ for $t_k \leq t \leq t_{k+1}$ with $e^{g(t)} = f(t)$ in $[t_0, t_1]$. It follows easily that $e^{g(t)} = f(t)$ $\forall t \in [a, b]$. The fact that if $c \in [a, b]$ and $f(c) = e^z$ then we can choose g such that $g(c) = z$ and the uniqueness of g are both obvious from the fact that $e^{z_1} = e^{z_2}$ if and only if $z_1 - z_2 = 2n\pi i$ for some integer n. [Note that if n depends on $t \in [a, b]$ and if $t \to n$ is continuous then n is necessarily a constant]. If the domain $[a, b]$ of f is replaced by $\mathbb{R}$, we can find continuous functions $g_n : [-n, n] \to \mathbb{C}$ such that $e^{g_n} = f$ on $[-n, n]$. Since $e^{g_n} = e^{g_n + 1}$ on $[-n, n]$ we have $g_{n+1} = g_n + 2k_n\pi i$ where k_n is an integer valued function on $[-n, n]$. By continuity, k_n is actually a constant, Replacing g_{n+1} by $g_{n+1} - 2k_n\pi i$ we can make sure that $g_{n+1} = g_n$ on $g_{n+1} = g_n + 2k_n\pi i$. An induction argument now shows that $g_n's$ can be defined consistently on $\mathbb{R}$. This gives us a continuous function g on $\mathbb{R}$ with $e^g = f$. Note, in particular, that g can be chosen to vanish at 0 if f is a characteristic function with no zeros.

Theorem 3

If ϕ is an i.d. characteristic function then $\phi(t) \neq 0$ $\forall t \in \mathbb{R}$.

Proof: for $n = 1, 2, \ldots$ let ϕ_n be a characteristic function such that $\phi_n^n = \phi$. Let $\psi(t) = |\phi(t)|^2$ and $\psi_n(t) = |\phi_n(t)|^2$ for $n = 1, 2, \ldots$. Then $\psi_n(t) = \psi^{1/n}(t)$. Hence $\lim_{n\to\infty} \psi_n(t) = \begin{cases} 1 \text{ if } \psi(t) \neq 0 \\ 0 \text{ if } \psi(t) = 0 \end{cases}$. Recall that ψ and ψ_n are characteristic functions. Since $\psi(t) \neq 0$ for $|t|$ sufficiently small it follows (by continuity theorem) that $\lim_{n\to\infty} \psi_n(t)$ is necessarily continuous. Hence $\lim_{n\to\infty} \psi_n(t) \equiv 1$ and $\psi(t) \neq 0$ $\forall t$. This implies $\phi(t) \neq 0$ $\forall t$.

Is the converse true? In other words, if ϕ is a characteristic function which never vanishes can we conclude that ϕ is i.d.? The answer is no. In fact if X takes the values $0, 1$ and -1 with probabilities $\frac{3}{4}, \frac{1}{8}$ and $\frac{1}{8}$ then $Ee^{itX} = \frac{3}{4} + \frac{1}{8}e^{it} + \frac{1}{8}e^{-it} = \frac{3}{4} + \frac{1}{4}\cos t$ which never vanishes. Of course, X is not i.d..

Combining above theorem and the lemma before it we conclude that if ϕ is an i.d. characteristic function then there is a unique continuous function g on $\mathbb{R}$ such that $g(0) = 0$ and $e^{g(t)} = \phi(t)$ $\forall t$. Throughout these notes we write $\log \phi$ for g. g is called the *distinguished logarithm* of ϕ. Note that if ϕ is a *non-negative* i.d. (hence strictly positive) characteristic function with $\phi = \phi_n^n$ where f_n is a characteristic function then $\phi_n = \phi^{1/n}$ because $(\frac{\phi_n}{\phi^{1/n}})^n \equiv 1$ so $\frac{\phi_n}{\phi^{1/n}}$ is an $n-$ the root of unity. However $\frac{\phi_n(t)}{\phi^{1/n}(t)}$ is continuous so $\frac{\phi_n}{\phi^{1/n}} = c$ where c is an $n-$ th root of unity independent of t. But $\phi_n(0) = \phi^{1/n}(0) = 1$

so $\phi_n \equiv \phi^{1/n}$. In particular $\phi^{1/n}$ is a characteristic function for each n. Also $g(t) = \log \phi(t)$ is the natural logarithm of the positive number $\phi(t)$.

Exercise

If f is a strictly positive characteristic function then $\sqrt{f}$ need not be a characteristic function

Hint: $\frac{3+\cos t}{4}$ is the characteristic function of a random variable taking values $0, 1$ and -1 with probabilities $\frac{3}{4}, \frac{1}{8}$ and $\frac{1}{8}$ respectively. It is impossible to find i.i.d. random variable X and Y such that $X+Y$ has this distribution. (Why?). Hence $\sqrt{\frac{2+\cos t}{3}}$ is not a characteristic function

Theorem 4

If $\{\phi_n\}$ is a sequence of i.d. characteristic functions and $\phi_n \to \phi$ pointwise where ϕ is continuous at 0 then ϕ is i.d..

Proof: the Continuity Theorem for characteristic functions shows that ϕ is indeed a characteristic function. We claim that $\phi(t) \neq 0 \ \forall t$. Recall that $|\phi|^2$ and $|\phi_n|^2$ are characteristic functions. Let $\psi(t) = |\phi(t)|^2$ and $\psi_n(t) = |\phi_n(t)|^2$. Then $\psi_n^{1/m}(t) \to \psi^{1/m}(t)$ for $m = 1, 2, \ldots$. As observed earlier $\psi_n^{1/m}$ is a characteristic function. By Continuity Theorem $\psi^{1/m}$ is also a characteristic function. It follows that ψ is i.d.. Hence ψ never vanishes implying that ϕ never vanishes. We have proved the claim. Now we recall that for a characteristic function ϕ with no zeros we defined $g = \log \phi$ by $g(t) = l(\phi(t))$ if $t \in [t_0, t_1]$ and $g(t) = g(t_k) + l(\frac{\phi(t)}{\phi(t_k)})$ if $t_k \leq t \leq t_{k+1}$ for $k = 1, 2, \ldots, N-1$ where $\{t_j\}$ is a suitable partition. [If $\rho = \inf |\phi(t)| : |t| \leq N\}$ and δ is chosen such that $|\phi(t) - \phi(s)| < \frac{\rho}{2}$ whenever $|t - s| \leq \delta$ the partition $\{t_j\}$ is chosen such that $t_{j+1} - t_j \leq \delta$]. Now let $g_n = \log \phi_n$ be defined by $g_n(t) = l(\phi_n(t))$ if $t \in [t_0, t_1]$ and $g_n(t) = g_n(t_k) + l(\frac{\phi_n(t)}{\phi_n(t_k)})$ if $t_k \leq t \leq t_{k+1}$ for $k = 1, 2, \ldots, N-1$; If $\rho_n = \inf |\phi_n(t)| : |t| \leq N\}$ and δ_n is chosen such that $|\phi_n(t) - \phi_n(s)| < \frac{\rho_n}{2}$ the partition $\{t_j\}$ can be chosen such that $t_{j+1} - t_j \leq \delta_n$: recall from basic theory of characteristic functions that $\phi_n \to \phi$ uniformly on $[-N, N]$. Ignoring a finite number of integers n we can find ρ_0 independent of n such that $0 < \rho_0 \leq \min\{|\phi_n(t)|, |\phi(t)|\}$. Also the condition $|\phi_n(t) - \phi_n(s)| < \frac{\rho}{2}$ whenever $|t - s| \leq \delta$ holds with δ independent of n (by uniform convergence). Hence the same partition can be used in the definitions of g and $g_n' s$. Since l is uniformly continuous on the compact set $\{z : |z - 1| \leq \frac{1}{2}\}$ it is clear that $g_n \to g$ uniformly on $[-N, N]$. If $m \in \mathbb{N}$ then $e^{g/m} = \lim e^{g_n/m}$ and $e^{g_n/m}$ is a characteristic function. (Why?). Applying the Continuity Theorem again we conclude that $e^{g/m}$ is a characteristic function. Since $(e^{g/m})^m = \phi$ and m is arbitrary we have proved that ϕ is i.d..

Notation: if μ is an i.d. probability measure on $\mathbb{R}$ and $n \geq 1$, there is a unique probability measure μ_n on $\mathbb{R}$ such that the $n-$fold convolution of μ_n with itself if μ. We write $\mu^{1/n}$ for μ_n.

We now discover more i.d. distributions using above theorem; our aim is to characterize all i.d. characteristic functions.

Theorem 5

Let ν be a finite positive Borel measure on $\mathbb{R}$ with $\nu(\{0\}) = 0$ and

$$\phi(t) = e^{ict} e^{\int\limits_{-\infty}^{\infty} (e^{itx} - 1 - \frac{itx}{1+x^2}) \frac{1+x^2}{x^2} d\nu(x)}$$

where $(e^{itx} - 1 - \frac{itx}{1+x^2}) \frac{1+x^2}{x^2}$ is defined as $-\frac{t^2}{2}$ at $x = 0$. Then ϕ is an i.d. characteristic function.

Remark: $-\frac{t^2}{2} = \lim\limits_{x \to 0} (e^{itx} - 1 - \frac{itx}{1+x^2}) \frac{1+x^2}{x^2}$.

Proof: it is enough to prove that ϕ is a characteristic function for any finite measure ν because we can get the $n-$ th root of ϕ by replacing ν by $\frac{1}{n}\nu$. We first observe that $e^{\lambda(e^{itx} - 1 - \frac{itx}{1+x^2})}$ is a characteristic function for any $\lambda > 0$ and $x \in \mathbb{R}$. In fact, if X has Poisson distribution with parameter λ then the characteristic function of $xX - \frac{\lambda x}{1+x^2}$ is $e^{\lambda(e^{itx} - 1 - \frac{itx}{1+x^2})}$. From this we claim that $e^{\int\limits_{-\infty}^{\infty} (e^{itx} - 1 - \frac{itx}{1+x^2}) d\tau(x)}$ is a characteristic function for any finite measure τ. Since this function is continuous it suffices show that it is a point-wise limit of characteristic functions. Now $\int\limits_{-\infty}^{\infty} (e^{itx} - 1 - \frac{itx}{1+x^2}) d\tau(x) = \lim\limits_{N \to \infty} \int\limits_{-N}^{N} (e^{itx} - 1 - \frac{itx}{1+x^2}) d\tau(x)$.

Hence it suffices to show that $e^{\int\limits_{a}^{b} (e^{itx} - 1 - \frac{itx}{1+x^2}) d\tau(x)}$ is a characteristic function for any $a < b$. If $\{x_j : 0 \leq j \leq k\}$ is the partition obtained by dividing $[a, b]$ into n equal parts then $\sum\limits_{j=1}^{k} (e^{itx_j} - 1 - \frac{itx_j}{1+x_j^2}) I_{[x_{j-1}, x_j)} - \sum\limits_{j=1}^{k} (e^{itx} - 1 - \frac{itx}{1+x^2}) \to 0$ as $n \to \infty$ (in fact uniformly on $[a, b]$) so it suffices to show that $e^{\lambda_j \{e^{itx_j} - 1 - \frac{itx_j}{1+x_j^2}\}}$ is a characteristic function, where $\lambda_j = \tau\{x_{j-1}, x_j)$. Since $e^{\lambda\{e^{itc} - 1 - \frac{itc}{1+c^2}\}}$ is a characteristic function for any $\lambda > 0$ and $c \in \mathbb{R}$, the proof of the claim is complete. Thus $e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})d\tau(x)}$ is a characteristic function for any finite measure τ. Now $e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu(x)} = \lim e^{\int_{\{|x|>\frac{1}{n}\}}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu(x)} = \lim e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})d\tau_n(x)}$ where $d\tau_n(x) = I_{\{|x|>\frac{1}{n}\}}\frac{1+x^2}{x^2}d\nu(x)$. Since each τ_n is a finite measure the proof of the theorem is complete.

Remark:

$\phi(t) = e^{-\sigma^2t^2/2}e^{ict}e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu(x)}$ is also an i.d. characteristic function if $\sigma > 0$.

The converse of this is also true: any i.d. characteristic function is of this type for some real number c, some $\sigma > 0$ and some finite measure ν.

Before proving this we prove that c, σ, ν are uniquely determined by $e^{ict}e^{-\sigma^2t^2/2}e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu(x)}$.

If $e^{ic_1t}e^{-\sigma_1^2t^2/2}e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_1(x)} = e^{ic_2t}e^{-\sigma_2^2t^2/2}e^{\int_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_2(x)}$ $\forall t$ then $c_1 = c_2, \sigma_1 = \sigma_2$ and $\nu_1 = \nu_2$. We first take absolute values and logarithms on both sides to get

$$-\sigma_1^2t^2/2+\int_{-\infty}^{\infty}(\cos(tx)-1)\frac{1+x^2}{x^2}d\nu_1(x) = -\sigma_2^2t^2/2+\int_{-\infty}^{\infty}(\cos(tx)-1)\frac{1+x^2}{x^2}d\nu_2(x).$$

Replacing t by t/a and multiplying by a^2 we get $-\sigma_1^2t^2/2 + a^2\int_{-\infty}^{\infty}(\cos(\frac{tx}{a}) - 1)\frac{1+x^2}{x^2}d\nu_1(x) = -\sigma_2^2t^2/2+a^2\int_{-\infty}^{\infty}(\cos(\frac{tx}{a})-1)\frac{1+x^2}{x^2}d\nu_2(x)$. We claim that $a^2\int_{-\infty}^{\infty}(\cos(\frac{tx}{a})-1)\frac{1+x^2}{x^2}d\nu_j(x) \to 0$ as $a \to 0$ for $j = 1, 2$. This would show that $\sigma_1 = \sigma_2$. Since $1 - \cos(\frac{tx}{a}) \leq \frac{t^2x^2}{a^2}$ the we have $a^2\left|(\cos(\frac{tx}{a}) - 1)\frac{1+x^2}{x^2}\right| \leq a^2\frac{t^2x^2}{a^2}\frac{1+x^2}{x^2} = t^2(1 + x^2)$. Hence, by Dominated Convergence Theorem, $a^2\int_{\{|x|\leq 1\}}(\cos(\frac{tx}{a}) -$

$1)\frac{1+x^2}{x^2}d\nu_j(x) \to 0$ as $a \to 0$ for $j = 1, 2$. Also $a^2 \int\limits_{\{|x|>1\}} (\cos(\frac{tx}{a})-1)\frac{1+x^2}{x^2}d\nu_j(x) \leq$ $2a^2 \int\limits_{\{|x|>1\}} \frac{1+x^2}{x^2}d\nu_j(x) \leq 4a^2\nu_j(\mathbb{R}) \to 0$ as $a \to 0$ for $j = 1, 2$. We have now proved that $\sigma_1 = \sigma_2$. Hence $e^{ic_1t}e^{\int\limits_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_1(x)} = e^{ic_2t}e^{\int\limits_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_2(x)}$ $\forall t$ which implies $ic_1t + \int\limits_{-\infty}^{\infty}(e^{itx} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_1(x) = ic_2t + \int\limits_{-\infty}^{\infty}(e^{itx} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_2(x)$ $\forall t$. [The two sides of this equation may differ by an integer multiple of 2π; however, the two sides are continuous and vanish at 0 and hence they are equal]. Replacing t be $t+s$, then by $t-s$. and adding the two equations we get

$2ic_1t+\int\limits_{-\infty}^{\infty}(2e^{itx}\cos(sx)-2-\frac{2itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_1(x) = 2ic_2t+\int\limits_{-\infty}^{\infty}(2e^{itx}\cos(sx) - 2 - \frac{2itx}{1+x^2})\frac{1+x^2}{x^2}d\nu_2(x)$ $\forall t, s$. Replacing t in the previous equation by $2t$ and subtracting the resulting equation from this equation we get

$\int\limits_{-\infty}^{\infty} e^{itx}(\cos(sx)-1)\frac{1+x^2}{x^2}d\nu_1(x) = \int\limits_{-\infty}^{\infty} e^{itx}(\cos(sx)-1)\frac{1+x^2}{x^2}d\nu_2(x)$ $\forall t, s$. It follows that the finite measures $(\cos(sx)-1)\frac{1+x^2}{x^2}d\nu_1(x)$ and $(\cos(sx)-1)\frac{1+x^2}{x^2}d\nu_2(x)$ have the same characteristic function. Hence, these two measures are equal. It follows that ν_1 and ν_2 coincide on Borel subsets of $\{x : \cos(sx) \neq 1\}$ for every $s \in \mathbb{R}$. Since there is no x such that $\cos(x) = 1$ and $\cos(\sqrt{2}x) = 1$ it follows that $\nu_1 = \nu_2$. (The reader is asked to fill in the details of this argument). It is now obvious that $c_1 = c_2$.

A modified version of this is the following:

$\int e^{itx}d\mu(x) = e^{ict}e^{-\sigma^2t^2/2}e^{\int\limits_{-\infty}^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})d\nu(x)}$ $\forall t$ for some real number c and positive number σ where ν is a positive measure such that $\int\limits_{-\infty}^{\infty}\frac{x^2}{1+x^2}d\nu(x) < \infty$. This representation is called the *Levy-Khinchine representation* and the unique measure ν is called the *Levy measure* of μ.

It would follow from the next theorem that every i.d. probability measure μ has a unique Levy measure ν related by an equation of the type

$\int e^{itx} d\mu(x) = e^{ict} e^{-\sigma^2 t^2/2} e^{\int_{-\infty}^{\infty} (e^{itx} - 1 - \frac{itx}{1+x^2}) d\nu(x)}$ $\forall t$ for some real number c and positive number σ. Further, there is always an i.d. measure μ corresponding to any positive measure ν such that $\int_{-\infty}^{\infty} \frac{x^2}{1+x^2} d\nu(x) < \infty$, any real number c and any positive number σ.

Theorem 6

Every i.d. characteristic function ϕ is of the type $\phi(t) = e^{ict} e^{-\sigma^2 t^2/2} e^{\int_{-\infty}^{\infty} (e^{itx} - 1 - \frac{itx}{1+x^2}) d\nu(x)}$ $\forall t$ for some real number c, some positive measure ν satisfying the condition $\int_{-\infty}^{\infty} \frac{x^2}{1+x^2} d\nu(x) < \infty$ and positive number σ.

Remark: the proof is somewhat lengthy and first time readers may read it at a later stage.

We need two lemmas:

Lemma 7

Let ϕ be i.d. and ϕ_n be the characteristic function whose $n-$th power is ϕ $(n = 1, 2, \ldots)$. Then $\limsup_{n\to\infty} n\mu_n\{x : |x| > a\} \leq a\alpha \int_0^{1/a} |\operatorname{Re} g(t)|\, dt$ where μ_n is the measure whose characteristic function is ϕ_n, $a > 0$ and $g = \log \phi$ and $\alpha = \frac{1}{\inf\{(1 - \frac{\sin t}{t} : |t| \geq 1\}}$.

[Recall that g is the unique continuous function such that $e^g = \phi$ and $g(0) = 0$].

To prove the lemma we start with the standard inequality $\mu_n\{x : |x| > a\} \leq a\alpha \int_0^{1/a} \{1 - \operatorname{Re} \phi_n(t)\} dt$. [Proposition 8.29, p. 171 of Probability by Breiman] where $\alpha = \frac{1}{\inf\{(1 - \frac{\sin t}{t} : |t| \geq 1\}}$. We claim that $n(\phi_n(t) - 1) \to g(t)$ $\forall t$. This follows from the fact that $e^{\frac{g(t)}{n}} = \phi_n(t)$ $\forall t$ and the fact that $n(e^{\frac{z}{n}} - 1) \to z$ $\forall z \in \mathbb{C}$. [To see that $e^{\frac{g}{n}} = \phi_n$ note that both sides have the same $n-$ th power; by continuity of g and ϕ_n it follows that $e^{\frac{g}{n}}/\phi_n$ is a constant which must be 1 because $g(0) = 0$ and $\phi_n(0) = 1$]. It follows now that $\limsup_{n\to\infty} n\mu_n\{x : |x| > a\} \leq$

$\limsup_{n\to\infty} a\alpha n \int_0^{1/a} \{1 - \operatorname{Re}\phi_n(t)\}dt = -a\alpha \int_0^{1/a} \operatorname{Re} g(t)dt \le a\alpha \int_0^{1/a} |\operatorname{Re} g(t)|\, dt$. [We have used Dominated Convergence Theorem here; note that $\{1 - \operatorname{Re}\phi_n(t)\} = \{1 - \operatorname{Re} e^{\frac{g(t)}{n}}\} \le \left|e^{\frac{g(t)}{n}} - 1\right| \le (e^{\frac{C}{n}} - 1)$ where $C = \sup\{|g(t)| : 0 \le t \le \frac{1}{a}\}$ and that $n(e^{\frac{C}{n}} - 1) \to C$ as $n \to \infty$ which implies that $\{n(e^{\frac{C}{n}} - 1)\}$ is bounded.

Lemma 8

$$\limsup_{n\to\infty} n \int_{\{x:|x|\le 1\}} x^2 d\mu_n(x) < \infty.$$

Proof of the lemma: $n\{1 - \operatorname{Re}\phi_n(t)\} \le n \int_{-\infty}^{\infty} (1 - \cos x\} d\mu_n(x) \ge n \int_{\{|x|\le 1\}} (1 - \cos x\} d\mu_n(x) \ge n\beta \int_{\{|x|\le 1\}} x^2 d\mu_n(x)$ where $\beta = \inf\{\frac{1-\cos x}{x^2} : |x| \le 1\}$. [We interpret $\frac{1-\cos x}{x^2}$ as $\frac{1}{2}$ when$x = 0$; note that $\beta > 0$]. It follows that $\limsup_{n\to\infty} n \int_{\{x:|x|\le 1\}} x^2 d\mu_n(x) \le \frac{1}{\beta} \limsup_{n\to\infty} n\{1 - \operatorname{Re}\phi_n(t)\} = -\frac{1}{\beta} \operatorname{Re} g(t) < \infty$.

Finally, we prove the Levy-Khinchine formula for ϕ.

Let λ_n be defined by $d\lambda_n(x) = \frac{nx^2}{1+x^2} d\mu_n(x)$. We have $\lambda_n(\mathbb{R}) = \int_{\{x:|x|\le 1\}} \frac{nx^2}{1+x^2} d\mu_n(x) + \int_{\{x:|x|>1\}} \frac{nx^2}{1+x^2} d\mu_n(x)$. By

Lemma 8, $\limsup \int_{\{x:|x|\le 1\}} \frac{nx^2}{1+x^2} d\mu_n(x) \le \limsup \int_{\{x:|x|\le 1\}} nx^2 d\mu_n(x) < \infty$. By Lemma 7, $\limsup \int_{\{x:|x|>1\}} \frac{nx^2}{1+x^2} d\mu_n(x) \le \limsup n\mu_n\{x : |x| > 1\} < \infty$. Hence $\{\lambda_n\}$ is a sequence of positive finite measures with $\sup\{\lambda_n(\mathbb{R}) : n \ge 1\} < \infty$. Let $\nu_n = \frac{\lambda_n}{\lambda_n(\mathbb{R})}$. We have $g(t) = \lim n\{\phi_n(t) - 1) = \lim \int_{-\infty}^{\infty} n(e^{ixt} - 1) d\mu_n(x)$

$$= \lim\{\int_{-\infty}^{\infty} (e^{ixt} - 1 - \tfrac{itx}{1+x^2}) \tfrac{1+x^2}{x^2} d\lambda_n(x) + int \int_{-\infty}^{\infty} \tfrac{x}{1+x^2} d\mu_n(x)\}$$

$= \lim\{\lambda_n(\mathbb{R}) \int\limits_{-\infty}^{\infty} (e^{ixt}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2} d\nu_n(x)+it\beta_n\}$ where $\beta_n = n \int\limits_{-\infty}^{\infty} \frac{x}{1+x^2} d\mu_n(x)$.

The rest of the proof is along the following lines: we show that $\liminf \lambda_n(\mathbb{R}) > 0$ and that $\{\nu_n\}$ is tight. It will then follow that for some integers $n_1 < n_2 < \ldots$, $\{\lambda_{n_j}\}$ converges to a positive number ρ and $\{\nu_{n_j}\}$ converges weakly to a probability measure ν_0. Since $(e^{ixt}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2}$ is a bounded continuous function on $\mathbb{R}$ it follows that $\beta = \lim \beta_{n_j}$ necessarily exists and $g(t) = \rho \int\limits_{-\infty}^{\infty} (e^{ixt} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2} d\nu_0(x) + it\beta$. Recall that $(e^{ixt} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2}$ equals $-\frac{t^2}{2}$ when $x = 0$. Let ν_1 be the restriction of $\rho\nu_0$ to $\mathbb{R}\backslash\{0\}$. Then we get

$$\phi(t) = e^{g(t)} = e^{\int\limits_{-\infty}^{\infty} (e^{ixt}-1-\frac{itx}{1+x^2})\frac{1+x^2}{x^2} d\nu_1(x)+it\beta-\frac{t^2}{2}\rho\nu_0\{0\}} = e^{\int\limits_{-\infty}^{\infty} (e^{ixt}-1-\frac{itx}{1+x^2}) d\nu(x)+it\beta-\frac{t^2}{2}\rho\nu_0\{0\}}$$

where $d\nu(x) = \frac{1+x^2}{x^2} d\nu_1(x)$. This finishes the proof.

Proof of $\rho \equiv \liminf \lambda_n(\mathbb{R}) > 0$: if $\liminf \lambda_n(\mathbb{R}) = 0$ then there exists $n_k \uparrow \infty$ such that $\lambda_{n_k}(\mathbb{R}) \to 0$. Hence $\int\limits_{-\infty}^{\infty} (e^{ixt} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2} d\lambda_{n_k}(x) \to 0$. [The integrand is a bounded function of x for fixed t]. But $g(t) = \lim\{\int\limits_{-\infty}^{\infty} (e^{ixt} - 1 - \frac{itx}{1+x^2})\frac{1+x^2}{x^2} d\lambda_{n_k}(x) + it\beta_{n_k}\}$ so $g(t) = \lim it\beta_{n_k}$ $\forall t$. This implies that $g(t) = it\beta$ for some β and hence $\phi(t) = e^{it\beta}$ in which case there is nothing to prove. It remains to show that $\{\nu_n\}$ is tight. We have $\nu_n\{|x| > a\} = \frac{\lambda_n\{|x|>a\}}{\lambda_n(\mathbb{R})} = \frac{1}{\lambda_n(\mathbb{R})} \int\limits_{\{|x|>a\}} \frac{nx^2}{1+x^2} d\mu_n \leq \frac{n\mu_n\{|x|>a\}}{\lambda_n(\mathbb{R})}$ and hence $\limsup \nu_n\{|x| > a\} \leq \frac{a\alpha}{\rho} \int\limits_{0}^{1/a} |\mathrm{Re}\, g(t)|\, dt$ by Lemma 7. Tightness of $\{\nu_n\}$ follows clear since $g(t) \to 0$ as $t \to 0+$.

Proposition 9

A positive measure ν is the Levy measure of an i.d. measure if and only if $\int \min\{1, x^2\} d\nu(x) < \infty$.

Proof: we only have to show that $\int \frac{x^2}{1+x^2} d\nu(x) < \infty$ if and only if $\int \min\{1, x^2\} d\nu(x) < \infty$. For this it suffices to observe that $\frac{x^2}{1+x^2} \leq \min\{1, x^2\}$ and $\min\{1, x^2\} \leq \frac{2x^2}{1+x^2}$.

Notation: we write $\mu[\beta,\sigma,\nu]$ for the probability measure μ whose characteristic function is $e^{\int_{-\infty}^{\infty}(e^{ixt}-1-\frac{itx}{1+x^2})d\nu(x)+it\beta-\frac{t^2\sigma^2}{2}}$.

Theorem 10

If $\mu[c_n,\sigma_n,\nu_n]\to\mu[c,\sigma,\nu]$ then $c_n\to c$ and $\nu_n|_{\{|x|>\delta\}}\to\nu|_{\{|x|>\delta\}}$ for every $\delta>0$. It does not follow that $\sigma_n\to\sigma$.

We do not prove this theorem here. A proof for i.d. laws on Banach spaces will be given in Volume 2 of these notes.

Definition: two probability measures μ and ν on $\mathbb{R}$ are said to be of the *same type* if there exist real numbers a and b with $a>0$ such that $\mu(E)=\nu(aE+b)$ for every Borel set E. Two random variables X and Y are of the same type if $Y\overset{d}{=}aX+b$ for some $a>0$ and $b\in\mathbb{R}$. This is equivalent to the fact that the induced measures are of the same type. Two characteristic functions ϕ_1 and ϕ_2 are said to be of the same type if $\phi_1(t)=e^{ibt}\phi_2(at)\ \forall t\in\mathbb{R}$. Random variables X and Y are of the same type if and only if their characteristic functions are.

Exercise

Verify that being of the same type is an equivalence relation in the class of probability measures or the class of characteristic functions. Also verify that the equivalence class of the standard normal distribution is precisely the class of all normal distributions.

Theorem 11 [Convergence of Types Theorem]

Let $X_n\overset{d}{\to}X$ and $a_nX_n+b_n\overset{d}{\to}Y$ where $a_n>0, b_n\in\mathbb{R}\ \forall n$. If X and Y are non-degenerate then there exist numbers $a>0$ and $b\in\mathbb{R}$ such that $a_n\to a, b_n\to b$ and $Y\overset{d}{=}aX+b$.

Proof: let ϕ_1,ϕ_2 and ψ_n be the characteristic functions of X,Y and X_n respectively. Then $\psi_n\to\phi_1$ and $e^{itb_n}\psi_n(a_nt)\to\phi_2(t)$ uniformly on compact sets. If $\{a_n\}$ is not bounded then some subsequence $\{a_{n_k}\}$ will increase to ∞. In this case, for any $\Delta>0$, $\sup\limits_{|t|\le\Delta}\left|e^{itb_{n_k}}\psi_{n_k}(a_{n_k}t)\to\phi_2(t)\right|\to 0$. For any $t\in\mathbb{R}$, $\frac{t}{a_{n_k}}\in[-\Delta,\Delta]$ for k sufficiently large so $e^{itb_{n_k}/a_{n_k}}\psi_{n_k}(t)-\phi_2(\frac{t}{a_{n_k}})\to 0$. This shows that $e^{itb_{n_k}/a_{n_k}}\psi_{n_k}(t)\to 1$. But $\psi_{n_k}(t)\to\phi_1(t)$ so $e^{itb_{n_k}/a_{n_k}}\to\frac{1}{\phi_1(t)}$ provided $|t|$ is so small that $\phi_1(t)\neq 0$. As a consequence of this $\lim\frac{b_{n_k}}{a_{n_k}}$ $(=c,$

say) exists and $\phi_1(t) = e^{-itc}$ for $|t|$ sufficiently small. This makes ϕ_1 generate which is a contradiction. We have proved that $\{a_n\}$ is bounded. Let a_0 be a limit point of $\{a_n\}$. Arguing as above we see that $\sup\limits_{|t|\leq\Delta} \left|e^{itb_{n_k}}\psi_{n_k}(a_{n_k}t) \to \phi_2(t)\right| \to 0$ and $\psi_{n_k}(a_{n_k}t) \to \phi_1(a_0 t)$ with $a_{n_k} \to a_0$. It follows that $e^{itb_{n_k}} \to \frac{\phi_2(t)}{\phi_1(a_0 t)}$ for $|t|$ sufficiently small; this implies $b = \lim b_{n_k}$ exists and $e^{itb}\phi_1(a_0 t) = \phi_2(t)$ for $|t|$ sufficiently small. Note that a_0 cannot be 0 because ϕ_2 is non-degenerate. If a_0' is another limit point of $\{a_n\}$ we get $|\phi_1(at)| = |\phi_1(a_0' t)|$ for $|t|$ sufficiently small. If $a_0' < a_0$ then $\left|\phi_1(\frac{a_0'}{a_0}t)\right| = |\phi_1(t)|$ for $|t|$ sufficiently small which implies $\left|\phi_1\{(\frac{a_0'}{a_0}\}^n t)\right| = |\phi_1(t)|$ $\forall n$ for $|t|$ sufficiently small and hence $|\phi_1(t)| = 1$ for $|t|$ sufficiently small. This is a contradiction. A similar argument shows that we cannot have $a_0 < a_0'$. Hence $\{a_n\}$ has a unique limit point a and $a > 0$. Now $\left|e^{itb_n}\phi_1(at) - e^{itb_n}\psi_n(a_n t)\right| = |\phi_1(at) - \psi_n(a_n t)| \to 0$ and $e^{itb_n}\psi_n(a_n t) \to \phi_2(t)$ so $e^{itb_n}\phi_1(at) \to \phi_2(t)$ $\forall t$. It follows that $\lim b_n = b$ exists and $e^{itb}\phi_1(at) = \phi_2(t)$ $\forall t$ which says $Y \overset{d}{=} aX + b$.

Exercise

Let ϕ be i.d. and $g = \log\phi$. If $c > 0$ show that e^{cg} is also an i.d. characteristic function.

Can you prove this without using the Levy-Khinchine Representation Theorem?

Theorem 12 [Schoenberg]

Let $f : \mathbb{R} \to \mathbb{C}$ be continuos with $f(0) = 0$. Then e^f is the characteristic function of an infinitely divisible distribution μ if and only if the following conditions hold:

a) $\bar{f}(-x) = f(x)$ $\forall x$

b) $\sum\limits_{j,k=1}^{N} c_j\bar{c}_k f(t_j - t_k) \geq 0$ whenever $N \in \mathbb{N}, t_j \in \mathbb{R}, c_j \in \mathbb{C}$ for $1 \leq j \leq N$ and $\sum\limits_{j=1}^{N} c_j = 0$.

Proof: if e^f is the characteristic function of an infinitely divisible distribution then $e^{\alpha f}$ is a characteristic function for each $\alpha > 0$. Hence it is positive definite for each $\alpha > 0$. But $\frac{e^{\alpha f}-1}{\alpha} \to f$ as $\alpha \to 0$ so $\sum\limits_{j,k=1}^{N} c_j\bar{c}_k f(t_j - t_k) =$

$\lim\limits_{\alpha\to 0}\sum\limits_{j,k=1}^{N} c_j\bar{c}_k \frac{e^{\alpha f(t_j-t_k)}-1}{\alpha} = \lim\limits_{\alpha\to 0}\sum\limits_{j,k=1}^{N} c_j\bar{c}_k \frac{e^{\alpha f(t_j-t_k)}}{\alpha} \geq 0$. Thus b) holds. a) follows from the fact that $e^{\bar{f}(-x)} = e^{f(x)}$ so $\bar{f}(-x)-f(x) = 2\pi i n(x)$ for some integer $n(x)$; by continuity and the fact that $f(0)=0$ we get $n(x)=0\ \forall x$ so a) holds. Conversely suppose a) and b) hold. Suppose $N\in\mathbb{N}, t_j\in\mathbb{R}, c_j\in\mathbb{C}$ for $1\leq j\leq N$. Let $c_0 = -\sum\limits_{j=1}^{N} c_j, t_0 = 0$. Then $\sum\limits_{j=0}^{N} c_j = 0$ and the hypothesis implies that $\sum\limits_{j,k=0}^{N} c_j\bar{c}_k f(t_j-t_k)\geq 0$. Hence $\sum\limits_{j,k=1}^{N} c_j\bar{c}_k f(t_j-t_k)+2\,\mathrm{Re}\{c_0\sum\limits_{j=1}^{N} c_j f(t_j)\}\geq 0$. We can rewrite this as $\sum\limits_{j,k=1}^{N} c_j\bar{c}_k\{f(t_j-t_k)-f(t_j)-f(-t_k)\}\geq 0$. Consider the $N\times N$ matrix $A = ((f(t_j-t_k)-f(t_j)-f(-t_k)))$. Then $A=A^*$. Also, A is positive definite. Hence there exists a unitary matrix C such that CAC^{-1} is a diagonal matrix. Let $d_1, d_2, \ldots, d_N$ be the diagonal entries. Then the entries of A are given by $a_{jk} = \sum\limits_{l=1}^{N}\bar{c}_{lj}\bar{c}_{lk}d_l$. Hence $f(t_j-t_k)-f(t_j)-f(-t_k) = \sum\limits_{l=1}^{N}\bar{c}_{lj}\bar{c}_{lk}d_l$. Let $S = \sum\limits_{j,k=1}^{N} c_j\bar{c}_k e^{\alpha f(t_j-t_k)} = \sum\limits_{j,k=1}^{N} b_j\bar{b}_k e^{\alpha[f(t_j-t_k)-f(t_j)-f(-t_k)]}$ where $b_j = c_j e^{\alpha f(t_j)}$. Hence $S = \sum\limits_{j,k=1}^{N} b_j\bar{b}_k e^{\alpha\sum\limits_{l=1}^{N}\bar{c}_{lj}\bar{c}_{lk}d_l}$. We now write $e^{\alpha\sum\limits_{l=1}^{N}\bar{c}_{lj}\bar{c}_{lk}d_l}$ as $\prod\limits_{l=1}^{N} e^{\alpha\bar{c}_{lj}\bar{c}_{lk}d_l}$, expand the exponentials and multiply out to get a sum of terms of the type $\left|\sum\limits_{j=1}^{N} b_j(\bar{c}_{lj})^{k_l}\right|^2 \alpha^{k_l}\frac{d_1^{k_1}d_2^{k_2}\ldots d_N^{k_N}}{k_1!k_2!\ldots k_N!}$ (each k_j varying from 0 to N). It follows that the continuous function $e^{\alpha f}$ is positive definite. Hence it is a characteristic function for each α and $e^{\frac{1}{n}f}$ is an $n-th$ root of e^f.

We now construct an importable class of i.d. distributions called *stable distributions*. These distributions have applications in Mathematical Finance and some other areas.

Definition: a Borel probability measure μ on $\mathbb{R}$ is called *stable* if $\mu^{*(n)}$ is of the same type as μ for each n where $\mu^{*(n)}$ is the convolution of μ with itself n times. Equivalently, $X_1+X_2+\ldots+X_n \overset{d}{=} a_nX_1+b_n, n=1,2,\ldots$ for some $\{a_n\}\subseteq(0,\infty)$ and some $\{b_n\}\subseteq\mathbb{R}$ where $\{X_i\}$ is i.i.d. with common distribution μ.

Remark: stability here is in the sense of stability with respect to i.i.d. sums. It is not related to any form of stability in the physical sense.

Examples

Constant random variables are stable. Here are some non-constant examples:

a) the normal characteristic function $e^{-t^2/2}$ is stable . Since $(e^{-t^2/2})^n = e^{-(\sqrt{n}t)^2/2}$ we can take $a_n = \sqrt{n}$ and $b_n = 0$.

b) the Cauchy characteristic function $e^{-|t|}$ is stable . Here $a_n = n, b_n = 0$.

c) Let X have the $N(0,1)$ distribution and $Y = \frac{1}{X^2}$ if $X \neq 0, 0$ if $X = 0$. Then Y has a stable distribution with $a_n = n^2$ and $b_n = 0$. (In particular this shows there are positive random variables with a stable distribution). This example requires a basic knowledge of Laplace transforms. To prove that Y is stable, we begin by observing that the density function of Y is given by

$f(x) = \frac{1}{\sqrt{2\pi}} x^{-3/2} e^{-1/(2x)} I_{(0,\infty)}(x)$. By the lemma below $\int\limits_0^\infty e^{-(a^2x^2+\frac{b^2}{x^2})} dx = \frac{\sqrt{\pi}}{2a} e^{-2ab}$ if $a, b \in (0,\infty)$. Now

$\int\limits_0^\infty e^{-tx} f(x) dx = \frac{1}{\sqrt{2\pi}} \int\limits_0^\infty e^{-tx} x^{-3/2} e^{-1/(2x)} dx = \frac{1}{\sqrt{2\pi}} \int\limits_0^\infty e^{-\frac{t}{y^2}} y^3 e^{-\frac{y^2}{2}} 2y^{-3} dy = \sqrt{\frac{2}{\pi}} \int\limits_0^\infty e^{-[\frac{y^2}{2}+\frac{t}{y^2}]} dy = e^{-\sqrt{2t}}$ for $t > 0$. If f_n is the $n-$ fold convolution of f with itself then $\int\limits_0^\infty e^{-tx} f_n(x) dx = (\int\limits_0^\infty e^{-tx} f(x) dx)^n = e^{-n\sqrt{2t}} = e^{-\sqrt{tn^2}}$ proving that f_n is the density function of $n^2 X$ where X is a random variable with density f. In the notations used above this means $Y_1 + Y_2 + \ldots + Y_n \stackrel{d}{=} n^2 Y_1$ if $\{Y_n\}$ is i.i.d. with the same distribution as Y.

Lemma 13

$$\int\limits_0^\infty e^{-(a^2x^2+\frac{b^2}{x^2})} dx = \frac{\sqrt{\pi}}{2a} e^{-2ab} \text{ if } a, b \in (0,\infty)$$

Proof of the lemma: let $I(a,b) = \int\limits_0^\infty e^{-(a^2x^2+\frac{b^2}{x^2})} dx$. Put $y = \sqrt{\frac{a}{b}} x$ to get

$$I(a,b) = \sqrt{\frac{b}{a}} \int\limits_0^\infty e^{-(aby^2+\frac{ab}{y^2})} dy$$

$= \sqrt{\frac{b}{a}} \int_0^\infty e^{-ab(y^2+\frac{1}{y^2})} dy = \sqrt{\frac{b}{a}} e^{-2ab} \int_0^\infty e^{-ab(y-\frac{1}{y})^2} dy$. Let $I = \int_0^\infty e^{-c(y-\frac{1}{y})^2} dy$ where $c = ab$. Put $x = \frac{1}{y}$ to get $I = \int_0^\infty e^{-c(x-\frac{1}{x})^2} \frac{1}{x^2} dx$. Adding these two equations we get $2I = I = \int_0^\infty e^{-c(x-\frac{1}{x})^2} [1 + \frac{1}{x^2}] dx = \int_0^\infty e^{-c(x-\frac{1}{x})^2} d(x - \frac{1}{x})$. Hence $2I = \int_0^\infty e^{-cu^2} du = \frac{1}{\sqrt{2c}} \int_0^\infty e^{-v^2/2} dv = \frac{\sqrt{2\pi}}{\sqrt{2c}}$. This gives $I = \frac{\sqrt{\pi}}{2\sqrt{ab}}$ and $I(a,b) = \sqrt{\frac{b}{a}} e^{-2ab} I = \sqrt{\frac{b}{a}} e^{-2ab} \frac{\sqrt{\pi}}{2\sqrt{ab}} = \frac{\sqrt{\pi}}{2a} e^{-2ab}$.

These two examples suggest that we could look at the functions $\phi(t) = e^{-c|t|^\alpha}$ where c and α are positive.

Exercise

Show that $e^{-c|t|^\alpha}$ is not a characteristic function if $\alpha > 2$.

Hint: consider the second derivative at 0.

Theorem 14

All stable distributions are i.d..

Proof: if $X_1 + X_2 + \ldots + X_n \stackrel{d}{=} a_n X_1 + b_n$ then $Y_1 + Y_2 + \ldots + Y_n \stackrel{d}{=} X_1$ where $Y_i = \frac{1}{a_n}(X_i - \frac{b_i}{n})$. Hence the distribution of the $X_i's$ has an n-th root for each n.

Now let $\phi(t) = e^{\int_{-\infty}^{\infty} (e^{ixt}-1-\frac{itx}{1+x^2}) d\nu(x)}$ where $d\nu(x) = \frac{1}{|x|^\alpha} dx$ for some $\alpha \in (1,3)$. Note that $\int \frac{x^2}{1+x^2} d\nu(x) < \infty$. Hence ν is a Levy measure and ϕ is an i.d. characteristic function. Since ν is symmetric it follows that $\mathrm{Im}(\int_{-\infty}^{\infty} (e^{ixt} - 1 - \frac{itx}{1+x^2}) d\nu(x)) = \int_{-\infty}^{\infty} (\sin(tx) - \frac{tx}{1+x^2}) d\nu(x) = 0$. Hence $\phi(t) = e^{\int_{-\infty}^{\infty} (\cos(tx)-1) d\nu(x)}$.

Consider $\int\limits_{-\infty}^{\infty}(\cos(tx)-1)d\nu(x) = \int\limits_{-\infty}^{\infty}(\cos(tx)-1)\frac{1}{|x|^\alpha}dx = -2\int\limits_{0}^{\infty}(1-\cos(tx))\frac{1}{x^\alpha}dx = -2(\int\limits_{0}^{\infty}(1-\cos(y))\frac{1}{y^\alpha}dy)\,|t|^{\alpha-1}$ by the substitution $y = x\,|t|$. Noting that $\int\limits_{0}^{\infty}(1-\cos(y))\frac{1}{y^\alpha}dy < \infty$ we have proved that $\phi(t) = e^{-c|t|^{\alpha-1}}$ for some $c > 0$. It follows that $e^{-c|t|^{\alpha-1}}$ is a characteristic function for some, hence for all, $c > 0$ provided $1 < \alpha < 3$. We have proved that $e^{-|t|^\alpha}$ is a characteristic function if $0 < \alpha < 2$. Since $(e^{-c|t|^\alpha})^n = e^{-c\left|n^{1/\alpha}t\right|^\alpha}$ these characteristic functions are stable.

It should be noted that $e^{-t^2/2}$ is a stable characteristic function and its Levy measure is the zero measure.

The stable measures we have constructed so far are all symmetric. We call $e^{-|t|^\alpha}$ an $SS(\alpha)$ characteristic function. $SS(\alpha)$ is read as "symmetric stable with index alpha".

We now consider properties of general stable distributions.

Proposition 15

Let μ be a non-degenerate stable probability measure. Then

a) the constants a_n, b_n are unique for each n.

b) if μ is symmetric then $b_n = 0\ \forall n$ but the converse is false

c) if $\tilde{\mu}(E) = \mu(-E)$ then $\mu * \tilde{\mu}$ is also stable with the same constants a_n and b_n.

Proof: suppose $a_nX_1 + b_n \overset{d}{=} c_nX_1 + d_n$. Then $X_1 \overset{d}{=} \frac{c_n}{a_n}X_1 + \frac{d_n - b_n}{a_n}$. If ϕ is the characteristic function of X_1 then $|\phi(t)| = \left|\phi(\frac{c_n}{a_n}t)\right|$. If $\frac{c_n}{a_n} < 1$ this gives $|\phi(t)| = \left|\phi((\frac{c_n}{a_n})^k t)\right|\ \forall k \in \mathbb{N}$ and $(\frac{c_n}{a_n})^k \to 0$ so $|\phi(t)| = 1\ \forall t$ which implies that ϕ is degenerate. If $\frac{c_n}{a_n} > 1$ we can replace t by $t\frac{a_n}{c_n}$ in $|\phi(t)| = \left|\phi(\frac{c_n}{a_n}t)\right|$ to get $|\phi(t)| = \left|\phi(\frac{a_n}{c_n}t)\right|$ and we conclude that ϕ is degenerate. Hence $a_n = c_n$. Now $X_1 \overset{d}{=} X_1 + \frac{d_n - b_n}{a_n}$ which implies $X_1 \overset{d}{=} X_1 + k\frac{d_n - b_n}{a_n}\ \forall k \in \mathbb{N}$. Clearly this implies $b_n = d_n$.

b) If μ is symmetric then $a_nX_1 + b_n \overset{d}{=} -a_nX_1 - b_n$ (because $X_1 + X_2 + \ldots + X_n$ is symmetric). Hence $X_1 + \frac{b_n}{a_n} \overset{d}{=} -X_1 - \frac{b_n}{a_n}$. Since X_1 is itself symmetric this gives $X_1 + \frac{b_n}{a_n} \overset{d}{=} X_1 - \frac{b_n}{a_n}$. This implies $X_1 \overset{d}{=} X_1 + k\frac{b_n}{a_n}$ for any positive integer

k. This is impossible unless $b_n = 0$ because $X_1 + k\frac{b_n}{a_n} \to \pm\infty$. Example c) above shows that b_n may be 0 for all n without μ being symmetric.

c) is obvious.

Theorem 16

Let μ be stable and non-degenerate. Then there exists a unique constant $\alpha \in (0, \infty)$ such that the constants a_n in the definition of stability are given by $a_n = n^{1/\alpha}$.

Proof: by part c) of previous theorem we may assume that μ is symmetric. Let $\{X_n\}$ be i.i.d. with distribution μ and $S_n = X_1 + X_2 + \ldots + X_n$ for $n = 1, 2 \ldots$. Then $S_{nk} = S_n + (X_{n+1} + X_{n+2} + \ldots + X_{2n}) + \ldots + (X_{(k-1)n+1} + X_{(k-1)n+2} + \ldots + X_{kn})$ (and the k terms on the right side are independent) which gives $a_{nk}X_1 \stackrel{d}{=} a_nX_1 + a_nX_2 + \ldots + a_nX_k$. Applying the definition of stability again this gives $a_{nk}X_1 \stackrel{d}{=} a_na_kX_1$. Since X_1 is non-constant it follows easily from this that $a_{nk} = a_na_k$ $\forall n, k \geq 1$. We now complete the proof by proving the following facts:

a) $\{a_n\}$ is increasing

b) $\frac{\log a_n}{\log n}$ is independent of $n \in \{2, 3, \ldots\}$

c) $a_n = n^{1/\alpha}$ with $\alpha \in (0, \infty)$

Note that c) is immediate from b). For a) we note that $X_1 > x$ implies $a_nX_1 + a_mX_2 > a_nx$ or $a_nX_1 - a_mX_2 > a_nx$. Hence $P\{X_1 > x\} \leq P\{a_nX_1 + a_mX_2 > a_nx\} + P\{a_nX_1 - a_mX_2 > a_nx\} = 2P\{a_nX_1 + a_mX_2 > a_nx\}$. Hence $P\{X_1 > x\} \leq 2P\{a_{n+m}X_1 > a_nx\}$ because $a_nX_1 + a_mX_2 \stackrel{d}{=} a_{n+m}X_1$. [This follows by writing S_{n+m} as $S_n + (X_{n+1} + X_{n+2} + \ldots + S_{n+m})$]. From this inequality we conclude that $\{\frac{a_n}{a_{n+m}} : n, m \geq 1\}$ is bounded. [If this sequence is not bounded we get $P\{X_1 > x\} = 0$ $\forall x > 0$. But μ is symmetric so $X_1 = 0$ a.s.]. Now $(\frac{a_n}{a_{n+1}})^k = \frac{a_{nk}}{a_{k(n+1)}}$ so $\{(\frac{a_n}{a_{n+1}})^k : k \geq 1\}$ is bounded implying that $\frac{a_n}{a_{n+1}} \leq 1$. This proves a).

If $j, k \geq 2$ and m is a positive integer there exists n_m such that $j^{n_m} \leq k^m < j^{n_m+1}$. [$n_m = [\frac{m \log k}{\log j}]$ where $[t]$ is the greatest integer not exceeding t]. Now we have $(a_j)^{n_m} \leq (a_k)^m \leq (a_j)^{n_m+1}$ so $n_m \log a_j \leq m \log a_k \leq (n_m + 1) \log a_j$. But $j^{n_m} \leq k^m < j^{n_m+1}$ so $n_m \log j \leq m \log k \leq (n_m + 1) \log j$. This gives $\frac{n_m \log a_j}{(n_m+1) \log j} \leq \frac{m \log a_k}{m \log k}$. Letting $m \to \infty$ and noting that $j^{n_m+1} > k^m \to \infty$ we see that $n_m \to \infty$ and $\frac{\log a_j}{\log j} \leq \frac{\log a_k}{\log k}$. But j and k are arbitrary integers greater than 1 so equality must hold in this last inequality. This proves b). The proof of the theorem is complete.

Definition: the number α in above theorem is called the *index* of stability. We say μ is $S(\alpha)$ or stable with index α.

Theorem 17

If μ is $S(\alpha)$ then $\int |x|^r \, d\mu(x) < \infty$ for $0 < r < \alpha$.

Remark: it can be shown that $\int |x|^r \, d\mu(x) = \infty$ if $r \geq \alpha$ provided $0 < \alpha < 2$. [For $\alpha = 2$ $\int |x|^r \, d\mu(x) < \infty$ for all $r > 0$ and there are no stable laws with $\alpha > 2$]. This will be proved later when we discuss Banach space valued stable random variables and their (so-called) spectral representations.

Proof: assume that μ is symmetric. See the exercise below for the general case. Let $\{X_1, X_2, \ldots, X_n\}$ be i.i.d with distribution μ. Then $P\{\max\{|X_i| : 1 \leq i \leq n\} > a\} \leq 2P\{|S_n| > a\}$. [This is well known inequality due to Paul Levy; we include a proof here for quick reference: let $E_j = \{|X_1| \leq a, |X_2| \leq a, \ldots |X_{j-1}| \leq a, |X_j| > a\}$ for $2 \leq j \leq n$ and $E_1 = \{|X_1| > a\}$. Let $T_n^{(j)} = -X_1 - X_2 - \ldots - X_{j-1} + X_j - X_{j+1} - \ldots - X_n$. Then $\frac{S_n + T_n^{(j)}}{2} = X_j$. Since $E_1, E_2, \ldots, E_n$ are disjoint events whose union is $\{\max\{|X_i| : 1 \leq i \leq n\} > a\}$ we get $P\{\max\{|X_i| : 1 \leq i \leq n\} > a\} = \sum_{j=1}^{n} P(E_j) \leq \sum_{j=1}^{n} P(E_j \cap \{|S_n| > a\}) + \sum_{j=1}^{n} P(E_j \cap \{\left|T_n^{(j)}\right| > a\})$

$= 2\sum_{j=1}^{n} P(E_j \cap \{|S_n| > a\}) \leq 2P\{|S_n| > a\}$]. Hence $P\{\max\{|X_i| : 1 \leq i \leq n\} > a\} \leq 2P\{n^{1/\alpha} |X_1| > a\}$. In other words, $1 - (P\{|X_1| \leq a\})^n \leq 2P\{n^{1/\alpha} |X_1| > a\}$. This implies that $e^{-nP\{|X_1| > a\}} \geq (1 - P\{|X_1| > a\})^n = P^n\{|X_1| \leq a\} \geq 1 - 2P\{n^{1/\alpha} |X_1| > a\}$. Changing a to $tn^{1/\alpha}$ we get $1 - e^{-nP\{|X_1| > tn^{1/\alpha}\}} \leq 2P\{|X_1| > t\}$ $\forall n$. We now use this this inequality to prove that $E|X_1|^r < \infty$ if $0 < r < \alpha$. We have $E|X_1|^r = \sum_{n=0}^{\infty} \int I_{\{tn^{1/\alpha} < |X_1| \leq t(n+1)^{1/\alpha}} |X_1|^r \, dP$

$\leq t^r \sum_{n=0}^{\infty} P\{tn^{1/\alpha} < |X_1| \leq t(n+1)^{1/\alpha}\}(n+1)^{r/\alpha} \leq t^r P\{|X_1| > 0\} + t^r \sum_{n=1}^{\infty} \{(n+1)^{r/\alpha} - n^{r/\alpha}\} P\{|X_1| > tn^{1/\alpha}\}$. Since $1 - e^{-nP\{|X_1| > tn^{1/\alpha}\}} \leq 2P\{|X_1| > t\}$ we get $nP\{|X_1| > tn^{1/\alpha}\} \leq \log \frac{1}{1 - 2P\{|X_1| > t\}}$ if t is so large that $2P\{|X_1| > t\} < 1$; hence $\sum_{n=1}^{\infty} \{(n+1)^{r/\alpha} - n^{r/\alpha}\} P\{|X_1| > tn^{1/\alpha}\} \leq \sum_{n=1}^{\infty} \frac{\{(n+1)^{r/\alpha} - n^{r/\alpha}\}}{n} C$ where $C = \log \frac{1}{1 - 2P\{|X_1| > t\}}$. The proof is complete since $\sum_{n=1}^{\infty} \frac{\{(n+1)^{r/\alpha} - n^{r/\alpha}\}}{n} < \infty$. [Apply Mean Value Theorem and note that $\sum_{n=1}^{\infty} \frac{(n+1)^{\frac{r}{\alpha}-1}}{n} < \infty$].

Exercise

Prove above result in the non-symmetric case.

Hint: $\mu * \tilde{\mu}$ is also stable with index α so $\int |x|^r (d\mu * \tilde{\mu})(x) < \infty$ for $0 < r < \alpha$. Use Fubini's Theorem to conclude that $\int |x|^r d\mu(x) < \infty$ for $0 < r < \alpha$. [Use the inequalities $(a+b)^p \leq c(a^p + b^p)$ $\forall a, b > 0$ where $c = 1$ if $0 < p < 1$ and $c = 2^{p-1}$ if $1 \leq p < \infty$].

Corollary 18

The index α of stability cannot exceed 2.

Proof: Assume that $\alpha > 2$ for some symmetric stable measure μ. By the theorem μ has finite variance σ^2. We have $S_n \stackrel{d}{=} n^{1/\alpha} X_1$ and so $\frac{S_n}{\sigma\sqrt{n}} \stackrel{d}{=} \frac{n^{1/\alpha}}{\sigma\sqrt{n}} X_1$. By the Central Limit Theorem $\frac{S_n}{\sigma\sqrt{n}}$ converges in distribution to the standard normal distribution. But $\frac{n^{1/\alpha}}{\sigma\sqrt{n}} X_1 \to 0$ since $\alpha > 2$.

Exercise

Prove without using above theorems that $e^{c|t|^\alpha}$ is not a characteristic function if $c > 0$ and $\alpha > 2$.

Hint: compute the second derivative at 0 and relate it to the second moment.

Theorem 19

Any non-degenerate $S(\alpha)$ measure μ is absolutely continuous with respect to Lebesgue measure.

Proof: let ϕ be the characteristic function of μ. If we show that $\phi \in L^1(\mathbb{R})$ it

would follow (by the inversion formula for characteristic functions) that $\mu << m$ (where m is the Lebesgue measure on $\mathbb{R}$). We claim that (in fact) $|\phi(t)| = e^{-c|t|^\alpha}$ for some $c > 0$. It is enough to show that $|\phi(t)|^2 = e^{-c|t|^\alpha}$ for some $c > 0$. Hence there is no loss of generality in assuming that $\phi \geq 0$. Since ϕ is infinitely divisible it never vanishes, so $\phi(t) > 0$ $\forall t$. It follows from stability that $\phi^n(t) = \phi(n^{1/\alpha} t)$. Hence $g = \log \phi$ satisfies the equation $ng(t) = g(n^{1/\alpha} t)$. The only continuous real functions satisfying this equation for all n and t are functions of the type $g(t) = c\,|t|^{\alpha}$. We ask the reader to supply proof. [Hint: let $h(t) = g(|t|^{1/\alpha})$. Then $h(nt) = nh(t)$ and h is continuous. Prove that $h(t) = c\,|t|]$. We now have $\phi(t) = e^{g(t)} = e^{c|t|^{\alpha}}$. Note that ϕ is bounded (and not identically 1) so $c < 0$. This finishes the proof.

Remark: we have proved above that all $SS(\alpha)$ characteristic functions are of the type $e^{-c|t|^{\alpha}}$. Can we find all $S(\alpha)$ characteristic functions? Yes, and this will be done a little later.

Remark: suppose μ is $S(\alpha)$ with $\alpha < 2$ (and non-degenerate). Then $|\phi(t)|^2 = e^{-c|t|^{\alpha}}$ for some $c > 0$. It follows easily that μ is not normal. Also note that $\alpha = 2$ implies $\frac{S_n}{\sqrt{n}} \overset{d}{=} X_1 + c_n\ \forall n$ (for some $\{c_n\} \subseteq \mathbb{R}$). We get $\frac{S_n - nEX_1}{\sqrt{n}\sqrt{Var(X_1)}} \overset{d}{=} \frac{1}{\sqrt{Var(X_1)}} X_1 + d_n$ for some $\{d_n\}$. Hence, by Central Limit Theorem we conclude that μ is normal. *Thus an $S(\alpha)$ is normal if and only if $\alpha = 2$.*

Definition: μ is *strictly stable* if the constants b_n in the definition of stability are all 0.

If $\alpha > 1$ then μ has finite mean. Let $\{X_n\}$ be i.i.d. with distribution μ. Then $X_1 + X_2 + \ldots + X_n \overset{d}{=} n^{1/\alpha} X_1 + b_n$ and $nEX_1 = n^{1/\alpha} EX_1 + b_n\ \forall n$. It follows that μ is strictly stable if and only if $EX_1 = 0$. For $\alpha < 1$ the notion is a bit more complicated.

Remark: it is easy to see that a normal distribution is strictly stable (with $\alpha = 2$) if and only if the mean is 0 if and only if the distribution is symmetric. If $1 < \alpha < 2$ and μ is strictly stable then the mean is 0 but the distribution need not be symmetric.

Theorem 20

If μ is $S(\alpha)$, non-degenerate and $\alpha \neq 1$ then there exists a unique $c \in \mathbb{R}$ such that $\mu * \delta_c$ is strictly stable with index α.

Remark: c is called the centering constant.

Proof: if $X_1 + X_2 + \ldots + X_n \overset{d}{=} n^{1/\alpha} X_1 + b_n$ and $Y_j = X_j + c$ then $Y_1 + Y_2 + \ldots + Y_n \overset{d}{=} n^{1/\alpha} Y_1 + d_n$ where $d_n = b_n + nc - cn^{1/\alpha}$. We begin by choosing the appropriate c for $n = 2$. Choose c to be $\frac{b_2}{2^{1/\alpha} - 2}$ so that $d_2 = 0$. Then $Y_1 + Y_2 \overset{d}{=} 2^{1/\alpha} Y_1$. It follows that $Y_1 + Y_2 + \ldots + Y_{2n} \overset{d}{=} 2^{1/\alpha}[Y_1 + Y_2 + \ldots + Y_n] \overset{d}{=} (2n)^{1/\alpha} Y_1 + 2^{1/\alpha} d_n$. On the other hand $Y_1 + Y_2 + \ldots + Y_{2n} \overset{d}{=} \{n^{1/\alpha} Y_1 + d_n\} + \{n^{1/\alpha} Y_2 + d_n\} \overset{d}{=} n^{1/\alpha} 2^{1/\alpha} Y_1 + 2d_n$. It follows that $2^{1/\alpha} d_n = 2d_n$ (since μ is not degenerate). This implies $d_n = 0$ so $Y_1 + Y_2 + \ldots + Y_n \overset{d}{=} n^{1/\alpha} Y_1$ for each n. Hence $\mu * \delta_c$, the distribution of $Y_1 = X_1 + c$ is strictly stable.

Remark: what this proof shows is that if $b_2 = 0$ then $b_n = 0\ \forall n$.

Remark: it is interesting to note that if μ is strictly $S(1)$ then so is $\mu * \delta_c$ for any c: if $X_1 + X_2 + \ldots + X_n \stackrel{d}{=} nX_1$ then $Y_1 + Y_2 + \ldots + Y_n \stackrel{d}{=} nY_1$ where $Y_j = X_j + c$. An example of an $S(1)$ measure which is not strictly stable is given later.

Exercise

Suppose the condition $X_1 + X_2 + \ldots + X_n \stackrel{d}{=} a_n X_1 + b_n$ in the definition of stability holds for $n = 2$. Can we conclude that the common distribution of $X_i's$ is stable?

Hint: no! Let $\nu = \sum\limits_{k=-\infty}^{\infty} \frac{1}{2^k}\delta_{2^k}$. Show that ν is a Levy measure. Let μ be i.d. with Levy measure ν.

Theorem 21

Let $\{X_n\}$ be i.i.d. random variables with distribution μ. If $X_1 + X_2 \stackrel{d}{=} aX_1$ and $X_1 + X_2 + X_3 \stackrel{d}{=} bX_1$ (where a and b are positive constants) then μ is stable.

Proof: assume that μ is not degenerate. Let ϕ be the characteristic function of μ. Then $\phi^2(t) = \phi(at)$ and $\phi^3(t) = \phi(bt)$. We first show that $\{\phi(t)\}^{2^n 3^m} = \phi(a^n b^m t)\ \forall n, m \in \mathbb{N},\ \forall t \in \mathbb{R}$. If this holds for a certain pair (n, m) $(\forall t)$ then $\phi(a^{n+1} b^m t) = \{\phi(a^n b^m t)\}^2 = \{\{\phi(t)\}^{2^n 3^m}\}^2 = \{\phi(t)\}^{2^{n+1} 3^m}$ so the equation holds for the pair $(n+1, m)$. Similarly, the equation also holds for $(n, m+1)$. Since the equation holds for $n = m = 1$ it holds for all n and m. Next, we show that ϕ never vanishes. Suppose, if possible, $\phi(t) = 0$. Then $\phi(a^n b^m t) = 0\ \forall n, m$. If $a < 1$ (or $b < 1$) then we get a contradiction be letting $n \to \infty$ (respectively $m \to \infty$). Now note that $\phi(t) = \{\phi(\frac{t}{a^n b^m})\}^{2^n 3^m}$. Hence $\phi(\frac{t}{a^n b^m}) = 0 \forall n, m$. This leads to a contradiction if $a > 1$ or $b > 1$. If $a = b = 1$ then it is easy to see that $X_1 = X_2 = 0$ a.s.. Thus, ϕ never vanishes. There exists a unique continuous function $g : \mathbb{R} \to \mathbb{C}$ such that $g(0) = 0$ and $e^{g(t)} = \phi(t)\ \forall t$. It follows easily from this that $g(a^n b^m t) = 2^n 3^m g(t)\ \forall n, m \in \mathbb{N},\ \forall t \in \mathbb{R}$. This equation holds if one (or both of) n, m is (are) 0. We claim that it holds for all integers n and m (positive, negative or 0). Suppose $n \geq 1$ and $m \geq 1$. Then $g(b^m t) = 2^n 3^m g(\frac{t}{a^n})$ so $3^m g(t) = 2^n 3^m g(\frac{t}{a^n})$. Hence $3^m g(b^m t) = 2^n 3^m g(\frac{b^m t}{a^n})$ or $3^{2m} g(t) = 2^n 3^m g(\frac{b^m t}{a^n})$ which says $g(a^{-n} b^m t) = 2^{-n} 3^m g(t)$. The remaining cases are similar. Thus, $g(a^n b^m t) = 2^n 3^m g(t)\ \forall n, m \in \mathbb{Z},\ \forall t \in \mathbb{R}$. Now consider $\{n \log a + m \log b : n, m \in \mathbb{Z}\}$. This is an additive subgroup of $\mathbb{R}$. Hence it is either dense of discrete. In the second case this subgroup is of the type $\{n\beta : n \in \mathbb{Z}\}$ for some $\beta > 0$. We claim that this case cannot occur. Assuming this claim for the moment we conclude that $\{n \log a + m \log b : n, m \in \mathbb{Z}\}$ is dense and this implies that $\{a^n b^m : n, m \in \mathbb{Z}\}$ is dense in $(0, \infty)$. Let $s > 0$. There exist sequences $\{n_j\}, \{m_j\}$ such that $a^{n_j} b^{m_j} \to s$. It follows that $\{2^{n_j} 3^{m_j}\}$ converges unless $g(t) = 0$. Let the limit be r. Then $g(st) = rg(t)$ provided $g(t) \neq 0$. Of course, r depends on s. Let $h(s) = r$. Then $g(st) = h(s)g(t)$ except when $g(t) = 0$. Note that $g(t) = 0, t \neq 0$ implies $\phi(t) = 1$ and hence $\phi(a^n b^m t) = 1 \ \forall n, m \in \mathbb{Z}$. This implies that $g \equiv 0$ and $\phi \equiv 1$. Thus, g never vanishes on $\mathbb{R}\backslash\{0\}$ and so the equation $g(st) = h(s)g(t)$ holds for all s and t. In particular $g(s) = h(s)g(1)$ so $h(s) = \frac{g(s)}{g(1)}$. Finally we get the functional equation $g(st)g(1) = g(s)g(t)$. This gives $h(st) = h(t)h(s), h(0) = 0$ and h is continuous. Since h is not identically zero this gives $h(t) = t^z$ for some complex number z for all $t > 0$. [My problem collection in real analysis contains a proof of this]. But $h(-t)h(-1) = h(t)$ so $h(-t) = \beta t^z$ where $\beta = \frac{1}{h(-1)}$. [$h(-1) = 0$ would imply $g(-1) = 0$ a contradiction]. We have proved that $\phi(t) = e^{g(1)t^z}$ if $t \geq 0$ and $\phi(t) = e^{\beta g(1)|t|^z}$ if $t < 0$. We now prove that $z \in \mathbb{R}$. Indeed, if $z = a + ib$ then there exists $t \in (0, \infty)$ such that $g(1)t^z = g(1)e^{a \log t} e^{ib \log t} = |g(1)| e^{a \log t}$. Since $|\phi| \leq 1$ we must have $\operatorname{Re}\{g(1)t^z\} \leq 0$ so we get $|g(1)| e^{a \log t} \leq 0$ which makes $g(1) = 0$ and ϕ degenerate. Hence z is real. It is clear now that $\phi^n(t) = \phi(n^{1/z} t)$ for all $t \in \mathbb{R}$. Hence ϕ is strictly stable. It remains to show that $\{n \log a + m \log b : n, m \in \mathbb{Z}\}$ cannot be of the form $\{n\beta : n \in \mathbb{Z}\}$ for any $\beta > 0$. Suppose this is the case. Then $\{a^n b^m : n, m \in \mathbb{Z}\} = \{c^k : k \in \mathbb{Z}\}$ where $c = e^\beta > 0$. There exist integers j, l such that $a = c^j$ and $b = c^l$. Now $g(a^n b^m t) = 2^n 3^m g(t)$ so $g(c^{jn} t) = g(a^n t) = 2^n g(t)$ and $g(c^{lm} t) = g(b^m t) = 3^m g(t)$. Taking $n = l$ and $m = j$ we see that $2^n g(t) = g(c^{jn} t) = g(c^{ml} t) = 3^m g(t)$. Taking $t \neq 0$ we get $2^n = 3^m$ which implies $n = m = 0$. Hence $a = c^j = c^n = 1$ and $b = c^l = c^n = 1$. This however leads to the contradiction $g(t) = 2^n 3^m g(t) \ \forall n, m, t$.

Remark: the following more general result is also true but we will not prove it here.

Theorem 22

Let $\{X_n\}$ be i.i.d. with distribution μ. If $X_1 + X_2 \stackrel{d}{=} aX_1 + c_1$ and $X_1 + X_2 + X_3 \stackrel{d}{=} bX_1 + c_2$ (where a and b are positive constants and c_1, c_2 are real numbers) then μ is stable.

In the next theorem $s(t) = 1$ if $t \geq 0, -1$ if $t < 0$.

Theorem 23 [A characterization of stability]

If $\alpha \neq 1$ then μ is $S(\alpha)$ if and only if for any a and $b > 0$ there exists $c > 0$ such that $aX + bY \stackrel{d}{=} (a^\alpha + b^\alpha)^{1/\alpha} X + c$ where X and Y are i.i.d. with distribution μ. If $\alpha = 1$ and μ is strictly $S(\alpha)$ then $aX + bY \stackrel{d}{=} (a + b)X + c$ where X and Y are i.i.d. with distribution μ.. The constant c vanishes if μ is strictly stable.

Remark: the corresponding result for non-symmetric μ with $\alpha = 1$ will be stated later as an exercise. [See the exercise immediately following Theorem 27].

Proof: the 'only if' part is trivial for $\alpha = 2$. Let $0 < \alpha < 2$ and μ be strictly stable. Then $S_{n+m} = S_n + (X_{n+1} + X_{n+2} + \ldots + X_{n+m})$ and hence $(n+m)^{1/\alpha}X_1 \stackrel{d}{=} n^{1/\alpha}X_1 + m^{1/\alpha}X_2$. If j, k, l, m are positive integers then $(\frac{k}{l})^{1/\alpha}X_1 + (\frac{j}{m})^{1/\alpha}X_2 = \frac{1}{(lm)^{1/\alpha}}\{(mk)^{1/\alpha}X_1 + (jl)^{1/\alpha}X_2\} \stackrel{d}{=} \frac{1}{(lm)^{1/\alpha}}(mk + jl)^{1/\alpha}X_1 = (\frac{k}{l} + \frac{j}{m})^{1/\alpha}X_1$. Letting $\frac{k}{l} \to a^\alpha$ and $\frac{j}{m} \to b^\alpha$ we get $aX_1 + bX_2 \stackrel{d}{=} (a^\alpha + b^\alpha)^{1/\alpha}X_1$. If $\alpha \neq 1$ then some translate of μ is strictly stable which implies that $aX + bY \stackrel{d}{=} (a^\alpha + b^\alpha)^{1/\alpha}X + c$ for some c. We now prove the converse. Suppose $\alpha \neq 1$ and, for every $a, b > 0$ there exists $c = c(a, b)$ with $aX + bY \stackrel{d}{=} (a^\alpha + b^\alpha)^{1/\alpha}X + c$. Then $X_1 + X_2 \stackrel{d}{=} 2^{1/\alpha}X_1 + c_1$ for some c_1. Hence $X_1 + X_2 + X_3 \stackrel{d}{=} 2^{1/\alpha}X_1 + c_1 + X_3 \stackrel{d}{=} (2+1)^{1/\alpha}X_1 + c_2$. An induction argument shows that $S_n \stackrel{d}{=} n^{1/\alpha}X_1 + c_n$ for some real number c_n. the proof is now complete.

Theorem 24 [Infinitely divisible distributions as limits of sums of independent random variables]

a) A probability measure μ is i.d. if and only if there exist random variables $\{X_{nj} : 1 \leq j \leq m_n, n = 1, 2, \ldots\}$ such that $\{X_{nj} : 1 \leq j \leq m_n\}$ is independent for each n, $\max\limits_{1 \leq j \leq m_n} P\{|X_{nj}| > \varepsilon\} \to 0$ as $n \to \infty$ for every $\varepsilon > 0$ and $\{\sum\limits_{j=1}^{m_n} X_{nj}\}$ converges in distribution to μ.

b) A probability measure μ is stable if and only if there exist i.i.d. random variables $X_n : n = 1, 2, \ldots$ such that $\frac{1}{a_n}\{\sum\limits_{j=1}^{n} X_j + b_n\}$ converges in distribution to μ for some sequences $\{a_n\}, \{b_n\} \subseteq \mathbb{R}$ with $a_n > 0\ \forall n$.

Remark: if $\frac{1}{a_n}\{\sum\limits_{j=1}^{n} X_j + b_n\}$ converges in distribution to μ for some sequences $\{a_n\}, \{b_n\} \subseteq \mathbb{R}$ with $a_n > 0\ \forall n$ where $\{X_n\}$ is i.i.d. with distribution ν we say ν is in the *domain of attraction* of μ. Part b) of the theorem says that μ has a domain of attraction (in the sense there is some measure ν in its domain of attraction) if and only if it is stable.

Proof: a) suppose μ is i.d.. For each n there exist i.i.d. random variables $X_{n1}, X_{n2}, .., X_{nn}$ such that the distribution of $X_{n1} + X_{n2} + ... + X_{nn}$ is μ. Let μ_n be the distribution of X_{n1}. Then $\max\limits_{1 \leq j \leq k_n} P\{|X_{nj}| > \varepsilon\} = \mu_n\{x : |x| > \varepsilon\} \to 0$ as $n \to \infty$ for every $\varepsilon > 0$ by the next lemma. Hence a) holds.

Conversely suppose there exist random variables $\{X_{nj} : 1 \le j \le m_n, n = 1, 2, \ldots\}$ such that $\{X_{nj} : 1 \le j \le m_n\}$ is independent for each n, $\max\limits_{1\le j\le m_n} P\{|X_{nj}| > \varepsilon\} \to 0$ as $n \to \infty$ for every $\varepsilon > 0$ and $\{\sum\limits_{j=1}^{m_n} X_{nj}\}$ converges in distribution to μ. We have $\left|1 - Ee^{itX_{nk}}\right| \le E\left|(1 - e^{itX_{nk}})I_{\{|X_{nk}|\le\varepsilon\}}\right| + E\left|(1 - e^{itX_{nk}})I_{\{|X_{nk}|>\varepsilon\}}\right| \le \varepsilon|t| + 2P\{|X_{nk}| > \varepsilon\}$. Hence $\max\limits_{1\le j\le m_n}\left|1 - Ee^{itX_{nk}}\right| \to 0$ as $n \to \infty$ uniformly on compact sets. Let $\tau(x) = \begin{cases} 1 \text{ if } x > 1 \\ -1 \text{ if } x < -1 \\ x \text{ if } -1 \le x \le 1 \end{cases}$. Let $t_{nk} (1 \le k \le m_n, n = 1, 2, \ldots)$ be chosen such that $E\tau(X_{nk} + t_{nk}) = 0$. This is possible because $E\tau(X_{nk} + t)$ is a continuous function of t, $E\tau(X_{nk} + t) \to 1$ as $t \to \infty$ and $E\tau(X_{nk} + t) \to -1$ as $t \to -\infty$. We claim that $\max\limits_{1\le j\le m_n} P\{|X_{nj} + t_{nk}| > \varepsilon\} \to 0$. From the fact that $E\tau(X_{nk} + t_{nk}) = 0$ it is clear that no subsequence $\{t_{n_jk_j}\}$ can tend to $\pm\infty$. In other words, the collection $\{t_{nk}\}$ is bounded. If $t = \lim t_{n_jk_j}$ and $t \ne 0$ then $X_{n_jk_j} \to 0$ in probability (by hypothesis) and hence $X_{n_jk_j} + t_{n_jk_j} \to t$ in probability. Hence $\tau(X_{n_jk_j} + t_{n_jk_j}) \to \tau(t)$ in probability. It follows that $0 = E\tau(X_{n_jk_j} + t_{n_jk_j}) \to \tau(t)$ implying that $\tau(t) = 0$, hence $t = 0$, contradiction. It follows that $\max\{|t_{nk}| : 1 \le k \le m_n\} \to 0$. From this and the hypothesis we get $\max\limits_{1\le j\le m_n} P\{|X_{nj} + t_{nk}| > \varepsilon\} \to 0$ for each $\varepsilon > 0$. The proof now reduces to the following: if $\{Y_{nj} : 1 \le j \le m_n, n = 1, 2, \ldots\}$ are such that $\{Y_{nj} : 1 \le j \le m_n\}$ is independent for each n, $\max\limits_{1\le j\le m_n} P\{|Y_{nj}| > \varepsilon\} \to 0$ as $n \to \infty$ for every $\varepsilon > 0, E\tau(Y_{nk}) = 0$ and $\{\sum\limits_{j=1}^{m_n} Y_{nj}\}$ converges in distribution to μ then μ is i.d.. Denoting the distribution of Y_{nk} by μ_{nk} we get $\left|1 - Ee^{itY_{nk}}\right| = \left|\int\{e^{itx} - 1 - it\tau(x)\}d\mu_{nk}\right| \le \frac{t^2}{2}\int\limits_{\{|x|\le1\}} x^2 d\mu_{nk} + \int\limits_{\{|x|>1\}} (2 + |t|)d\mu_{nk}$. Hence $\left|1 - Ee^{itY_{nk}}\right| \le \frac{t^2}{2}\int\limits_{\{|x|\le1\}}\{\tau(x)\}^2 d\mu_{nk} + \int\limits_{\{|x|>1\}}(2 + |t|)\{\tau(x)\}^2 d\mu_{nk} \le (\frac{t^2}{2} + 2 + |t|)E\{\tau(Y_{nj})\}^2$. Writing Log for the principle branch of logarithm (defined on $\mathbb{C}\backslash(-\infty, 0]$) we see that for any $\Delta > 0$, $LogEe^{itY_{nk}}$ is well defined for $|t| \le \Delta$ for all k provided n is sufficiently large. [This is because $\max\limits_{1\le k\le m_n}\left|1 - Ee^{itY_{nk}}\right| \to 0$ as $n \to \infty$ uniformly on compact sets]. Now $\left|LogEe^{-itY_{nk}} - \{Ee^{-itY_{nk}} - 1\}\right| \le 2\left|Ee^{-itY_{nk}} - 1\right|^2$ for $|t| \le \Delta$ and n sufficiently large. [We used the inequality $|Log(1 + z) - z| \le 2|z|^2$ for $|z| < \frac{1}{2}$: $|Log(1 + z) - z| = \left|-\frac{z^2}{2} + \frac{z^3}{3} - \ldots\right| \le \frac{|z|^2}{1-|z|} \le 2|z|^2$]. Also $\sum\limits_{k=1}^{m_n}\left|1 - Ee^{itY_{nk}}\right|^2 \le \max\limits_{1\le k\le m_n}\left|1 - Ee^{itX_{nk}}\right|\sum\limits_{k=1}^{m_n}\left|1 - Ee^{itY_{nk}}\right| \le \max\limits_{1\le k\le m_n}\left|1 - Ee^{itX_{nk}}\right|(\frac{t^2}{2} + 2 + |t|)c_n$

where $c_n = \sum_{j=1}^{m_n} E\{\tau(Y_{nj})\}^2$. It follows that $\sum_{k=1}^{m_n} Log E e^{-itY_{nk}} = \sum_{k=1}^{m_n}\{E e^{-itY_{nk}} - 1\} + o(c_n)$ uniformly for $|t| \leq \Delta$. Hence $\frac{1}{2\delta}\sum_{k=1}^{m_n}\int_{-\delta}^{\delta} Log E e^{itY_{nk}} dt = \frac{1}{2\delta}\sum_{k=1}^{m_n}\int_{-\delta}^{\delta}\{E e^{itY_{nk}} - 1\} dt + 0(c_n)$. We now claim that $\sum_{k=1}^{m_n} Log E e^{-itY_{nk}} - \sum_{k=1}^{m_n}\{E e^{itY_{nk}} - 1\} \to 0$ uniformly for $|t| \leq \Delta$. Since $E e^{it\sum_{k=1}^{m_n} Y_{nk}} \to \int e^{itx} d\mu(x)$ uniformly on compact sets we see that $\prod_{k=1}^{m_n} E e^{itY_{nk}} \to \int e^{itx} d\mu(x)$ and hence $\sum_{k=1}^{m_n} Log E e^{itY_{nk}} \to Log(\int e^{itx} d\mu(x))$ uniformly on compact sets. Hence $\frac{1}{2\delta}\sum_{k=1}^{m_n}\int_{-\delta}^{\delta} Log E e^{itY_{nk}} dt \to \frac{1}{2\alpha}\int_{-\delta}^{\delta} Log(\int e^{itx} d\mu(x))$. We have proved that $\frac{1}{2\delta}\sum_{k=1}^{m_n}\int_{-\delta}^{\delta}\{E e^{itY_{nk}} - 1\} dt + 0(c_n) \to \frac{1}{2\alpha}\int_{-\delta}^{\delta} Log(\int e^{itx} d\mu(x))$. But the left side here is $E\sum_{k=1}^{m_n}\{\frac{\sin(\delta Y_{nk})}{\delta Y_{nk}} - 1\} + 0(c_n)$. It is easy to verify that $1 - \frac{\sin x}{x} \geq a\{\tau(x)\}^2$ for some $a > 0$. It follows now that $aE\sum_{k=1}^{m_n}\{\tau(\delta Y_{nk})\}^2 + 0(c_n)$ remains bounded as $n \to \infty$. Since $\frac{\tau(x)}{\tau(\delta x)}$ is bounded we see that $a_0 c_n + 0(c_n)$ is bounded for some $a_0 > 0$. This implies that $\{c_n\}$ is bounded. Since $\sum_{k=1}^{m_n} Log E e^{-itY_{nk}} = \sum_{k=1}^{m_n}\{E e^{-itY_{nk}} - 1\} + o(c_n)$ we now see that $\sum_{k=1}^{m_n} Log E e^{-itY_{nk}} - \sum_{k=1}^{m_n}\{E e^{-itY_{nk}} - 1\} \to 0$ $\forall t$. Hence $\int e^{itx} d\mu(x) = \lim e^{\sum_{k=1}^{m_n} Log E e^{-itY_{nk}}} = \lim e^{\sum_{k=1}^{m_n}\{E e^{-itY_{nk}} - 1\}} = \lim e^{\int\{e^{itx} - 1) d\mu_n}$ where $\mu_n = \sum_{k=1}^{m_n} \mu_{nk}$. μ_{nk} being the distribution of Y_{nk}. We have expressed $\int e^{itx} d\mu(x)$ as the pointwise limit of a sequence of i.d. characteristic functions. [Indeed $e^{\int\{e^{itx} - 1) d\nu}$ is an i.d. characteristic function for any finite measure ν]. This completes the proof.

Proof of b): if μ is stable and $X_1, X_2, \ldots$ are i.i.d. random variables with distribution μ then and constants a_n, b_n with $a_n > 0$ such that $\frac{1}{a_n}\{\sum_{k=1}^{n} X_k + b_n\}$ has distribution μ for every n. In particular $\frac{1}{a_n}\{\sum_{k=1}^{n} X_k + b_n\}$ converges in distribution to μ. Conversely, suppose $\{X_n\}$ is i.i.d., $a_n > 0, b_n \in \mathbb{R}$ and $\frac{1}{a_n}\{\sum_{k=1}^{n} X_k + b_n\}$converges in distribution to μ. Let $Z_n = \frac{1}{a_n}\{S_n + b_n\}$ where $S_n = \sum_{k=1}^{n} X_k$. Fix a positive integer x and consider $\{Z_k, Z_{2k}, Z_{3k}, \ldots\}$. We can write Z_{nk} as $\frac{1}{a_{nk}}\{S_n^{(1)} + S_n^{(2)} + \ldots + S_n^{(k)}\} + \frac{b_{nk}}{a_{nk}}$ where $S_n^{(j)} = X_{(j-1)n+1} + X_{(j-1)n+2} + \ldots + X_{jn}$. Hence $\frac{1}{a_n}(S_n^{(1)} + b_n) + \frac{1}{a_n}(S_n^{(2)} + b_n) + \ldots + \frac{1}{a_n}(S_n^{(k)} + b_n) = \frac{a_{nk}}{a_n} Z_{nk} + \frac{b_{nk}}{a_n} - kb_n$. We conclude that $\frac{a_{nk}}{a_n} Z_{nk} + \frac{b_{nk}}{a_n} - kb_n$ converges in distribution to $\mu^{*(k)}$ (the $k-$ fold convolution of μ with itself); also Z_n converges to μ in distribution. The Convergence of Types Theorem [Theorem 11 above] shows that $\sum_{j=1}^{k} X_j \stackrel{d}{=} \alpha_k X_1 + \beta_k$ for some $\alpha_k > 0, \beta_k \in \mathbb{R}$ which completes the proof.

Lemma 25

If μ is i.d. and μ_n is the probability measure satisfying the equation $\mu_n * \mu_n * \ldots * \mu_n$ (n factors) $= \mu$ then $\mu_n \to \delta_0$ weakly.

Proof of the lemma: if ϕ_n is the characteristic function of μ_n and ϕ that of μ then we claim that $\log \phi_n(t) = \frac{\log \phi(t)}{n}$. For this note that $(e^{\frac{\log \phi(t)}{n}})^n = \phi(t) = (\phi_n(t))^n$ so $\frac{1}{\phi_n(t)} e^{\frac{\log \phi(t)}{n}}$ is an $n-$ th root of unity which is necessarily a constant (by continuity). Since this function has the value 1 at 0 we must have $\frac{1}{\phi_n(t)} e^{\frac{\log \phi(t)}{n}} = 1 \; \forall t$. Now the facts that $\log \phi_n(t)$ and $\frac{\log \phi(t)}{n}$ are both continuous, vanish at 0 and $e^{\frac{\log \phi(t)}{n}} = e^{\log \phi_n(t)} \equiv \phi_n(t)$ imply that $\log \phi_n(t) = \frac{\log \phi(t)}{n} \; \forall t$. It follows now that $\phi_n(t) = e^{\frac{\log \phi(t)}{n}} \to 1$ as $n \to \infty \; \forall t$. Hence $\mu_n \to \delta_0$ weakly.

Theorem 26

If μ is a symmetric $S(\alpha)$ probability measure with $1 < \alpha \leq 2$ then the support S of μ is $\{0\}$ or $\mathbb{R}$.

Proof: recall that $aX+bY \stackrel{d}{=} (a^\alpha+b^\alpha)^{1/\alpha} X$ if $\{X, Y\}$ is i.i.d with distribution μ and $a, b > 0$. It follows that $aS+bS = (a^\alpha+b^\alpha)^{1/\alpha} S$. Take $a = b = 1$ and note that $-S = S$ to conclude that $0 \in 2^{1/\alpha} S$. Hence $0 \in S$. Now taking $a^\alpha + b^\alpha = 1$ we get $aS \subseteq aS + bS \subseteq S$. Clearly this implies that $aS \subseteq S$ for $0 \leq a \leq 1$. Also taking $a = b = 2^{-1/\alpha}$ we get $2^{-1/\alpha} S + 2^{-1/\alpha} S = S$; hence $2^{1-1/\alpha} S \subseteq S$

(: if $x \in S$ then $2^{1-1/\alpha}x = 2^{-1/\alpha}x + 2^{-1/\alpha}x \in 2^{-1/\alpha}S + 2^{-1/\alpha}S = S$). By iteration we get $2^{n(1-1/\alpha)}S \subseteq S\ \forall n$. Together with the fact that $aS \subseteq S$ for $0 \leq a \leq 1$ this shows that $aS \subseteq S\ \forall a > 0$. This implies $aS = S\ \forall a > 0$. By symmetry the same equation holds for all $a \in \mathbb{R}$. Now we can go back to $aS + bS = (a^\alpha + b^\alpha)^{1/\alpha}S$ to conclude that $S + S = S$. Hence S is a subspace of $\mathbb{R}$. Of course, $\{0\}$ and $\mathbb{R}$ are the only subspaces of $\mathbb{R}$.

Remark: this proof works for stable measures on Banach spaces.

Proof: if μ is a measure then $\mu * \delta_x$ is a symmetric Gaussian measure for some x (namely $x = \int\limits_B y d\mu(y)$) and we can use the there with $\alpha = 2$.

Theorem 27 [Stable Characteristic Function]

Let μ be $S(\alpha)$ with characteristic function ϕ. Assume that $0 < \alpha < 2$.

a) If $\alpha \neq 1$ then $\log \phi(t) = -c\,|t|^\alpha \{\cos \frac{\pi\alpha}{2} + i\beta s(t)\} + idt$ where $c > 0$ or $c < 0$ according as $\alpha < 1$ or $\alpha > 1$, $\beta \in [-1, 1]$ and $d \in \mathbb{R}$.

b) If $\alpha = 1$ then $\log \phi(t) = -c\,|t|\,\{\frac{\pi}{2} + i\beta s(t) \log|t|\} + idt$ where $c > 0, d \in \mathbb{R}$ and $\beta \in [-1, 1]$.

Remark: conversely **above expressions are necessarily characteristic functions of stable distributions**. This result will not be proved here. See the remark immediately after the statement of Levy's Spectral Representation Theorem in Volume II)

Using uniqueness of Levy - Khinchine representation it is fairly straightforward to see that the Levy measure ν of an $S(\alpha)$ measure μ satisfies the relation $n\nu(E) = \nu(n^{-1/\alpha}E)$ or $n\nu(n^{1/\alpha}E) = \nu(E)$ for all Borel sets E. Note that ν is not the zero measure because μ is not normal. Let $h_1(x) = \int\limits_{(o,x]} y^2 d\nu(y), h_2(x) = \int\limits_{[-x,0)} y^2 d\nu(y)$ and $h(x) = h_1(x) + h_2(x) = \int\limits_{[-x,x]} y^2 d\nu(y)$ for $x > 0$. To find out how these functions look like we need two lemmas.

Lemma 28

Let $f : (0, \infty) \to [0, \infty)$ be a non-decreasing function such that $h(x) \equiv \lim\limits_{t\to\infty} \frac{f(tx)}{f(t)}$ exists and if finite $\forall x \in (0, \infty)$.

Then, either there exists $\rho \geq 0$ such that $h(x) = x^\rho\ \forall x$ or $h(x) = 0\ \forall x$.

Proof of the lemma: since $\frac{f(txy)}{f(t)} = \frac{f(txy)}{f(tx)}\frac{f(tx)}{f(t)}$ we have $h(xy) = h(y)h(x)$. Hence, if $h(x) = 0$ for some x then $h \equiv 0$. Assume now that $h(x) > 0\ \forall x$. Then $\log h(e^x)$ is an additive measurable function on $\mathbb{R}$ and hence there is a constant c such that $\log h(e^x) = cx$. This gives $h(x) = e^{\rho \log x} = x^\rho$ for some real number ρ. Since f is non-decreasing it follows that $\rho \geq 0$.

Lemma 29

Let $f : (0, \infty) \to [0, \infty)$ be a non-decreasing function such that $h(x) = \lim\limits_{n\to\infty} \beta_n f(a_n x)$ exists and $\in (0, \infty)\ \forall x$ where $\{a_n\}$ and $\{\beta_n\}$ are sequences of positive numbers such that $a_n \uparrow \infty$ and $\frac{\beta_{n+1}}{\beta_n} \to 1$. If h is also continuous then $h(x) = cx^\rho$ for some $c > 0$ and some $\rho \geq 0$.

If $t > a_1$ then there exists n such that $a_n \leq t < a_{n+1.}$ Hence $\frac{\beta_{n+1}}{\beta_n}\frac{\beta_n f(a_n x)}{\beta_{n+1} f(a_{n+1})} \leq \frac{f(tx)}{f(t)} \leq \frac{\beta_n}{\beta_{n+1}}\frac{\beta_{n+1} f(a_{n+1}x)}{\beta_n f(a_n)}$ which gives $\frac{h(x)}{h(1)} \leq \liminf\limits_{t\to\infty} \frac{f(tx)}{f(t)} \leq \limsup\limits_{t\to\infty} \frac{f(tx)}{f(t)} \leq \frac{h(x)}{h(1)}$ proving that $\lim\limits_{t\to\infty} \frac{f(tx)}{f(t)} = \frac{h(x)}{h(1)}$. By Lemma 29 there exists $\rho \geq 0$ such that $\frac{h(x)}{h(1)} = x^\rho$. The proof is complete.

Now let us recall the definitions $h_1(x) = \int\limits_{(o,x]} y^2 d\nu(y), h_2(x) = \int\limits_{[-x,0)} y^2 d\nu(y)$ and $h(x) = h_1(x) + h_2(x) = \int\limits_{[-x,x]} y^2 d\nu(y)$ for $x > 0$. Observe that $h(x) < \infty\ \forall x$ because $\int\limits_{\{|x|\leq 1\}} y^2 d\nu(y) < \infty$ and $\nu(\{x : |x| > 1\}) < \infty$. Continuity of h follows from the fact that ν is a continuous measure; indeed $n\nu\{x\} = \nu\{n^{-1/\alpha}x\}$ so $\sum\limits_{n=1}^{\infty} \frac{1}{n}\nu\{x\} = \sum\limits_{n=1}^{\infty} \nu\{n^{1/\alpha}x\} \leq \nu\{y : |y| \geq x\} < \infty$ forcing $\nu(x)$ to be 0. From the fact that $n\nu(E) = \nu(n^{-1/\alpha}E)$ we get $h(n^{1/\alpha}x) = n^{(\frac{2}{\alpha}-1)} \int\limits_{[-x,x]} z^2 d\nu(z) = n^{(\frac{2}{\alpha}-1)}h(x)$. Hence $n^{(1-\frac{2}{\alpha})}h(n^{1/\alpha}x) = h(x)$. By Lemma 30 it follows that $h(x) = cx^\rho$ for some $c > 0$ and some $\rho \geq 0$. By similar arguments the functions h_1 and h_2 also have the same form, say $h_j(x) = c_j x^{\rho_j}, j = 1, 2$. The equation $\int\limits_{(o,x]} y^2 d\nu(y) = c_1 x^{\rho_1}\ \forall x > 0$ implies that $x^2 d\nu(x) = c_1\rho_1 x^{\rho_1 - 1}dx$ or $d\nu(x) = c_1\rho_1 x^{\rho_1-3}dx$ on $(0, \infty)$. Similarly, $d\nu(x) = c_2\rho_2 |x|^{\rho_2-3} dx$ on $(-\infty, 0)$. We leave it as an exercise to show that $\rho_1 = \rho_2 = \rho$ unless $c_1 = 0$ or $c_2 = 0$. [Just look at the highest of the numbers ρ, ρ_1, ρ_2]. If $c_j = 0$ then ρ_j can be replaced by $\rho, j =$

$1, 2$. Hence we can assume that $\rho_1 = \rho_2 = \rho$. Then $d\nu(x) = c_1 \rho x^{\rho-3} dx$ on $(0, \infty)$ and $d\nu(x) = c_2 \rho |x|^{\rho-3} dx$ on $(-\infty, 0)$. To find the value of ρ we use the relation $\int \min\{1, y^2\} d\nu(y) < \infty$. This gives $0 < \rho < 2$. Now $n\nu(1, \infty) = \nu(n^{-1/\alpha}, \infty)$ so $\int_1^\infty n c_1 \rho x^{\rho-3} dx = \int_{n^{-1/\alpha}}^\infty c_1 \rho x^{\rho-3} dx$ or $n c_1 \rho \frac{1}{2-\rho} = c_1 \rho \frac{1}{\rho-2} n^{-\frac{1}{\alpha}(\rho-2)}$ which implies $\rho = 2 - \alpha$. We have shown that $d\nu(x) = \begin{cases} C_1 x^{-1-\alpha} dx \text{ on } (0, \infty) \\ C_2 |x|^{-1-\alpha} dx \text{ on } (-\infty, 0) \end{cases}$ (where $C_1 = c_1 \rho$ and $C_2 = c_2 \rho$). Let $\tau_\alpha(x) = \begin{cases} x \text{ if } \alpha > 1 \\ \sin x \text{ if } \alpha = 1 \\ 0 \text{ if } \alpha < 1 \end{cases}$.

Then $\int_{-\infty}^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| d\nu(x) = C_1 \int_0^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| x^{-1-\alpha} dx + C_2 \int_{-\infty}^0 \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| |x|^{-1-\alpha} dx$ $= C_1 \int_0^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| x^{-1-\alpha} dx + C_2 \int_0^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| x^{-1-\alpha} dx$. It is clear that both the terms on the right are finite if $0 < \alpha < 1$. If $\alpha > 1$ then $\int_0^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| x^{-1-\alpha} dx = \int_0^\infty \left|x - \frac{x}{1+x^2}\right| x^{-1-\alpha} dx = \int_0^\infty \frac{x^{2-\alpha}}{1+x^2} dx < \infty$ and, similarly, $\int_0^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| x^{-1-\alpha} dx < \infty$. if $\alpha = 1$ then $C_1 \int_0^\infty \left|\sin x - \frac{x}{1+x^2}\right| x^{-2} dx +$ $C_2 \int_0^\infty \left|\sin x - \frac{x}{1+x^2}\right| x^{-2} dx < \infty$ because $\frac{|\sin x - x|}{x^2}$ is bounded in $(0, 1)$. Thus $\int_{-\infty}^\infty \left|\tau_\alpha(x) - \frac{x}{1+x^2}\right| d\nu(x) < \infty$ in all cases. Using the Levy-Khinchine representation we can now write $\log \phi(t) = ict - t^2\sigma^2/2 + \int_{-\infty}^\infty \{e^{itx} - 1 - \frac{itx}{1+x^2}\} d\nu(x) =$ $idt - t^2\sigma^2/2 + \int_{-\infty}^\infty \{e^{itx} - 1 - it\tau_\alpha(x)\} d\nu(x)$ where $d = c - \int_{-\infty}^\infty \{\frac{x}{1+x^2} - \tau_\alpha(x)\} d\nu(x)$ where ν is the Levy measure of μ. Let $\psi(t) = \int_{-\infty}^\infty \{e^{itx} - 1 - it\tau_\alpha(x)\} d\nu(x)$. Recall that $n\nu(E) = \nu(n^{-1/\alpha} E)$. We can assert now that $t\nu(E) = \nu(t^{-1/\alpha} E)$ $\forall t > 0$. Indeed, this is an easy consequence of the fact that $d\nu(x) = \begin{cases} C_1 x^{-1-\alpha} dx \text{ on } (0, \infty) \\ C_2 |x|^{-1-\alpha} dx \text{ on } (-\infty, 0) \end{cases}$.

If $\alpha > 1$ then $\psi(t) = \int\limits_{-\infty}^{\infty} \{e^{itx} - 1 - it\tau_\alpha(x)\} d\nu(x) = \int\limits_{-\infty}^{\infty} \{e^{itx} - 1 - itx\} d\nu(x)$

$= \int\limits_{-\infty}^{\infty} \{e^{iy} - 1 - iy\} t^\alpha d\nu(y) = at^\alpha$ for $t > 0$, where $a = \int\limits_{-\infty}^{\infty} \{e^{iy} - 1 - iy\} d\nu(y)$.

Similarly, $\psi(t) = bt^\alpha$ where $b = \int\limits_{-\infty}^{\infty} \{e^{iy} - 1\} d\nu(y)$. We now evaluate a and b.

Let $\alpha < 1$. Then $b = C_1 \int\limits_0^{\infty} \{e^{iy} - 1\} y^{-1-\alpha} dx + C_2 \int\limits_{-\infty}^{0} \{e^{iy} - 1\} |y|^{-1-\alpha} dy$. To compute $\int\limits_0^{\infty} \{e^{iy} - 1\} y^{-1-\alpha} dy$ we consider $\int\limits_0^{\infty} \{e^{iy-\lambda y} - 1\} y^{-1-\alpha} dy$ where $\lambda > 0$. We have $\int\limits_0^{\infty} \{e^{iy-\lambda y} - 1\} y^{-1-\alpha} dy = \frac{y^{-\alpha}}{-\alpha} \{e^{iy-\lambda y} - 1\}|_0^\infty - \int\limits_0^{\infty} \frac{y^{-\alpha}}{-\alpha} \{e^{iy-\lambda y}(i - \lambda)\} dy =$ $\frac{i-\lambda}{\alpha} \int\limits_0^{\infty} y^{-\alpha} e^{iy-\lambda y} dy = \frac{i-\lambda}{\alpha} \lambda^{\alpha-1} \int\limits_0^{\infty} x^{-\alpha} e^{ix/\lambda - x} dx = \frac{i-\lambda}{\alpha} \lambda^{\alpha-1} \frac{\Gamma(1-\alpha)}{(1-\frac{i}{\lambda})^{1-\alpha}}$ using the fact that $\int\limits_0^{\infty} e^{itx} e^x x^{r-1} dx = \Gamma(r) \frac{1}{(1-t)^r}$ [This is the Gamma characteristic function]. Letting $\lambda \downarrow 0$ we get $\int\limits_0^{\infty} \{e^{iy} - 1\} y^{-1-\alpha} dx = \frac{i}{\alpha} e^{\pi i(1-\alpha)/2} \Gamma(1 - \alpha) =$ $-\frac{1}{\alpha} e^{-\pi i\alpha/2} \Gamma(1 - \alpha)$. Since $\int\limits_{-\infty}^{0} \{e^{iy} - 1\} |y|^{-1-\alpha} dy = \int\limits_0^{\infty} \{e^{-iy} - 1\} y^{-1-\alpha} dy$ is the complex conjugate of $\int\limits_0^{\infty} \{e^{iy} - 1\} y^{-1-\alpha} dx$ we get $b = -C_1 \frac{1}{\alpha} e^{-\pi i\alpha/2} \Gamma(1 -$ $\alpha) - C_2 \frac{1}{\alpha} e^{pi\alpha/2} \Gamma(1 - \alpha)$. Hence, $\log \phi(t) = idt - t^2\sigma^2/2 + \psi(t)$

$= idt - t^2\sigma^2/2 - C_1 \frac{1}{\alpha} t^\alpha e^{-\pi i\alpha/2} \Gamma(1 - \alpha) - C_2 t^\alpha \frac{1}{\alpha} e^{pi\alpha/2} \Gamma(1 - \alpha)$. Recall $\sigma = 0$ whenever $\alpha \neq 2$. Thus,

$\phi(t) = e^{idt} e^{-C_1 \frac{1}{\alpha} e^{-\pi i\alpha/2} \Gamma(1-\alpha) t^\alpha - C_2 \frac{1}{\alpha} e^{pi\alpha/2} \Gamma(1-\alpha) t^\alpha}$ for $t > 0$ and $0 < \alpha < 1$. This implies that $\phi(t) = e^{idt} e^{-|t|^\alpha \{C_1 e^{-\pi i\alpha/2} + C_2 e^{pi\alpha/2}\} \frac{1}{\alpha} \Gamma(1-\alpha)}$ $\forall t \in \mathbb{R}$ if $0 < \alpha <$ 1. From this we get $\phi(t) = e^{idt} e^{-|t|^\alpha \frac{\Gamma(1-\alpha)}{\alpha} (C_1 + C_2) \{\cos \frac{\pi\alpha}{2} + is(t) \frac{C_2 - C_1}{C_2 + C_1} \sin \frac{\pi\alpha}{2}\}}$. For $\alpha > 1$ we have to evaluate $\int\limits_0^{\infty} \{e^{iy-\lambda y} - 1 - iy\} y^{-1-\alpha} dy$. To do this we just note that $\frac{d}{dt}\int\limits_0^\infty \{e^{ity}-1-ity\}y^{-1-\alpha}dy = \int\limits_0^\infty \{e^{ity}-1\}(iy)y^{-1-\alpha}dy = i\int\limits_0^\infty \{e^{ity}-1\}y^{-\alpha}dy$ and the integral here has been computed. We leave it to the reader to complete the proof of the theorem for the case $\alpha > 1$.

Now let $\alpha = 1$. Recall that $\phi(t) = e^{ict-t^2\sigma^2/2+h(t)}$ where $h(t) = \int\limits_{-\infty}^\infty (e^{itx} - 1 - it\sin x)d\nu(x)$. Also, $d\nu(x) = c_1\frac{1}{x^2}I_{(0,\infty)} + c_2\frac{1}{x^2}I_{(-\infty,0)}$ and $\sigma = 0$ (as seen before). Hence $h(t) = c_1\int\limits_0^\infty \frac{e^{itx}-1-it\sin x}{x^2}dx + c_2\int\limits_{-\infty}^0 \frac{e^{itx}-1-it\sin x}{x^2}dx$. Consider $\mathrm{Re}\int\limits_0^\infty \frac{e^{itx}-1-it\sin x}{x^2}dx = \int\limits_0^\infty \frac{\cos tx-1}{x^2}dx = |t|\int\limits_0^\infty \frac{\cos y-1}{y^2}dy = -\frac{\pi}{2}|t|$. Also, $\mathrm{Im}\int\limits_0^\infty \frac{e^{itx}-1-it\sin x}{x^2}dx = \int\limits_0^\infty \frac{\sin(tx)-t\sin x}{x^2}dx$. Let $\varepsilon > 0, t > 0$ and consider $\int\limits_\varepsilon^\infty \frac{\sin(tx)-t\sin x}{x^2}dx = \int\limits_\varepsilon^\infty \frac{\sin(tx)}{x^2}dx - t\int\limits_\varepsilon^\infty \frac{\sin x}{x^2}dx = t\int\limits_{\varepsilon t}^\infty \frac{\sin(y)}{y^2}dy - t\int\limits_\varepsilon^\infty \frac{\sin x}{x^2}dx = t\int\limits_{\varepsilon t}^\varepsilon \frac{\sin x}{x^2}dx \to t\log\frac{1}{t}$ as $\varepsilon \to 0$. [We used the fact that $\frac{\sin x}{x} \to$ as $x \to 0+$ and $t\int\limits_{\varepsilon t}^\varepsilon \frac{1}{x}dx = -t\log t$].

Making obvious changes when $t < 0$ we get $\mathrm{Im}\int\limits_0^\infty \frac{e^{itx}-1-it\sin x}{x^2}dx = -t\log\frac{1}{|t|}$.

Thus $\int\limits_0^\infty \frac{e^{itx}-1-it\sin x}{x^2}dx = -\frac{\pi}{2}|t| - it\log\frac{1}{|t|}$. Now $h(t) = c_1\{-\frac{\pi}{2}|t| - it\log\frac{1}{|t|}\} + c_2\{-\frac{\pi}{2}|t| + it\log\frac{1}{|t|}\}$. [Because $\int\limits_{-\infty}^0 \frac{e^{itx}-1-it\sin x}{x^2}dx$ is the complex conjugate of $\int\limits_0^\infty \frac{e^{itx}-1-it\sin x}{x^2}dx$]. Thus, $\phi(t) = e^{ict+c_1\{-\frac{\pi}{2}|t|-it\log\frac{1}{|t|}\}+c_2\{-\frac{\pi}{2}|t|+it\log\frac{1}{|t|}\}}$.To complete the proof we write $c_1\{-\frac{\pi}{2}|t| - it\log\frac{1}{|t|}\} + c_2\{-\frac{\pi}{2}|t| + it\log\frac{1}{|t|}\}$ as $-|t|(c_1+c_2)\frac{\pi}{2} + it(c_2-c_1)\log\frac{1}{|t|} = -|t|\{(c_1+c_2)\frac{\pi}{2} + i(c_1-c_2)s(t)\log\frac{1}{|t|}\}$ where $s(t) = \begin{cases} 1 \text{ if } t > 0 \\ -1 \text{ if } t < 0 \\ 0 \text{ if } t = 0 \end{cases}$.

Writing C for $c_1 + c_2$ and β for $\frac{c_1-c_2}{c_1+c_2}$ we get $\phi(t) = e^{ict}e^{-C|t|\{\frac{\pi}{2}+i\beta s(t)\log|t|\}}$.

Exercise: suppose μ is $S(1)$, not necessarily strictly stable. Let X, Y be i.i.d with distribution μ and $a, b > 0$. Show that $aX + bY \stackrel{d}{=} (a+b)X + C\beta\{(a+b)\log(a+b) - a\log a - b\log b\}$ where $|\beta| \leq 1$ and $C > 0$. Also show that there exist $S(1)$ distributions which are not strictly stable (and no translates of which are strictly stable).

Hint: $e^{-c|t|\{\frac{\pi}{2}+i\beta s(t)\log|t|\}+idt}$ is always a stable characteristic function with $\alpha = 1$ if $c > 0, d \in \mathbb{R}$ and $\beta \in [-1, 1]$. [Take this for granted]. Take $\beta \neq 0$. In particular, $\phi(t) = e^{-|t|-\frac{2it}{\pi}\log|t|}$ $(c = \frac{2}{\pi}, \beta = 1)$; in this case $\phi^2(t) = \phi(2t)e^{\frac{4i}{\pi}(\log 2)t}$ or $X_1 + X_2 \stackrel{d}{=} 2X_1 + \frac{4}{\pi}\log 2$.

It can be shown that the support of a stable law is of the form $[a, \infty)$ or the form $(-\infty, a]$. If X is $N(0,1)$ then the support of $Y = \frac{1}{X^2}$ is $[0, \infty)$. [Indeed, $P\{a < Y < b\} > 0$ whenever $0 < a < b < \infty$]. Hence any interval of the type $[a, \infty)$ or of the type $(-\infty, a]$ is the support of a stable law. [Just look at translates of Y and $-Y$].

We state without proof a construction of stable random variables using exponential random variables.

Theorem 30

Let $0 < \alpha < 2$ and $\{\xi_i\}$ be i.i.d. random variables with $E\,|\xi_1|^\alpha < \infty$. Let $T_n = Y_1 + Y_2 + \ldots + Y_n$ where $\{Y_j\}$ is i.i.d. random variables with $P\{Y_1 \leq t\} = \begin{cases} 1 - e^{-t} \text{ if } t \geq 0 \\ 0 \text{ if } t < 0 \end{cases}$. Then the series $\sum_{n=1}^{\infty} T_n^{-1/\alpha}\xi_n$ converges a.s. and its sum is a stable random variable with index α.

Stable versus normal

If X is a normal random variable then $P\{X > t\}$ tends to 0 at an exponential rate as $t \to \infty$. In contrast, if X is stable with index $\alpha < 2$ then $t^\alpha P\{X > t\}$ converges to a positive finite limit. In other words, the tail probability $P\{X > t\}$ tends to 0 at the same rate as $t^{-\alpha}$. In view of this stable random variables are said to have a *heavy tail*. Stable distributions are used extensively in Mathematical Finance because of this heavy tail property. More results on stable laws appear on Volume 2 where infinitely divisible and stable laws on Banach spaces are discussed.

Positive Stable Distributions

Theorem 31

A probability measure μ supported by $[0,\infty)$ is stable if and only if $\int\limits_0^\infty e^{-tx}d\mu(x) = e^{-ct^\alpha - bt}$ $(t \geq 0)$ for some $c \geq 0, b \geq 0$ and $0 < \alpha < 1$.

We prove several propositions before proving above theorem.

Proposition 32

If X is a positive non-constant stable random variable then the index α of X is less than 1.

Proof: suppose, if possible, $\alpha = 1$. Let $\{X_n\}$ and S_n be as before and note that $S_2 \stackrel{d}{=} 2X_1 + b$ for some b. If $b > 0$ then $S_2 \geq b$ and hence $X_1 \geq \frac{b}{2}$ a.s.. But then $S_2 \stackrel{d}{=} 2X_1 + b \geq 2b$. This in turn implies $X_1 \geq b$ and hence $S_2 \geq 3b$, etc. By induction we get $S_2 \geq nb$ for every n, a contradiction. If $b < 0$ then $2X_1 + b \stackrel{d}{=} S_2$ so $2X_1 + b \geq 0$ and $X_1 \geq \frac{|b|}{2}$. This gives $S_2 \geq |b|$ and hence $2X_1 + b \stackrel{d}{=} S_2 \geq |b|$ which implies $X_1 \geq |b|$. This gives $S_2 \geq 2\,|b|$ etc. By induction we get $S_2 \geq n\,|b|$ for every n, a contradiction. We have proved that $b = 0$ and hence $S_2 \stackrel{d}{=} 2X_1$. This implies that if N is a positive integer then $S_{2^n} \stackrel{d}{=} 2^n X_1$. But $\frac{1}{2^n}S_{2^n} \geq \frac{1}{2^n}\sum\limits_{j=1}^{2^n}\min\{X_j, N\} \to E\min\{X_1, N\}$ a.s., by Strong Law of Large Numbers. Hence $X_1 \geq E\min\{X_1, N\}$ a.s. for every N. We have proved that $X_1 \geq EX_1$ a.s. which implies that EX_1 is finite and since $X_1 - EX_1$ is a non-negative random variable with zero mean, X_1 is degenerate. We have proved that $\alpha \neq 1$. Suppose $\alpha > 1$. Then $S_n \stackrel{d}{=} n^{1/\alpha}X_1 + b_n$ for some b_n. Also $EX_1 < \infty$. Hence $nEX_1 = ES_n = n^{1/\alpha}EX_1 + b_n$ and $b_n = (n - n^{1/\alpha})EX_1$. If $Y_n = X_n - EX_n$ then $Y_1 + Y_2 + \ldots + Y_n = S_n - nEX_1 \stackrel{d}{=} n^{1/\alpha}X_1 + b_n - nEX_1 = (n^{1/\alpha}Y_1 + n^{1/\alpha}EX_1) + b_n - nEX_1 = n^{1/\alpha}Y_1 \geq -n^{1/\alpha}EX_1$. Since $\{Y_j\}$ is i.i.d this implies $Y_1 \geq -\frac{1}{n}n^{1/\alpha}Y_1 \to 0$. Hence $X_1 \geq EX_1$ a.s. which forces X_1 to be degenerate.

Proposition 33

For every $\alpha \in (0,1)$ there exists a positive random variable with an $S(\alpha)$ distribution.

We claim that if $c > 0$ then $\frac{d^n}{dt^n}e^{-ct^\alpha} \geq 0$ if n is even and ≤ 0 if n is odd. For this we begin with $\frac{d}{dt}e^{-ct^\alpha} = -c\alpha t^{\alpha-1}e^{-ct^\alpha}$ and apply Leibniz rule to get $\frac{d^{n+1}}{dt^{n+1}}e^{-ct^\alpha} = \sum\limits_{k=0}^{n}\begin{pmatrix} n \\ k \end{pmatrix}\frac{d^k}{dt^k}e^{-ct^\alpha}\frac{d^{n-k}}{dt^{n-k}}(-c\alpha t^{\alpha-1})$. Note that $\frac{d^j}{dt^j}(-c\alpha t^{\alpha-1})$ is positive or negative according as j is odd or even. A simple induction argument proves our claim. [For instance, if we know that

$\frac{d^k}{dt^k}e^{-ct^\alpha} \geq 0$ if k is even and ≤ 0 if k is odd provided $k \leq n$ then $\frac{d^{n+1}}{dt^{n+1}}e^{-ct^\alpha} = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \frac{d^k}{dt^k}e^{-ct^\alpha} \frac{d^{n-k}}{dt^{n-k}}(-c\alpha t^{\alpha-1})$; suppose n is even. Then for k odd $\begin{pmatrix} n \\ k \end{pmatrix} \frac{d^k}{dt^k}e^{-ct^\alpha} \frac{d^{n-k}}{dt^{n-k}}(-c\alpha t^{\alpha-1}) \leq 0$ because $\frac{d^k}{dt^k}e^{-ct^\alpha} \leq 0$ and $\frac{d^{n-k}}{dt^{n-k}}(-c\alpha t^{\alpha-1}) \geq 0$. For k even $\frac{d^k}{dt^k}e^{-ct^\alpha} \geq 0$ and $\frac{d^{n-k}}{dt^{n-k}}(-c\alpha t^{\alpha-1}) \leq 0$ so we again have $\begin{pmatrix} n \\ k \end{pmatrix} \frac{d^k}{dt^k}e^{-ct^\alpha} \frac{d^{n-k}}{dt^{n-k}}(-c\alpha t^{\alpha-1}) \leq 0$. It follows that $\frac{d^{n+1}}{dt^{n+1}}e^{-ct^\alpha} \leq 0 \; \forall t$. Similar argument works for n odd and the claim follows by induction]. Now, by a well-known theorem on Laplace transforms (see, e.g. XIII.4, Theorem 1 of An Introduction to Probability Theory and its Applications by Willaim Feller, Vol. 2) we conclude that there exists a Borel probability measure μ on $[0, \infty)$ with $\int\limits_{[0,\infty)} e^{-tx} d\mu(x) = e^{-ct^\alpha}$ $(t \geq 0)$. Since the Laplace transform of the convolution of two measures in the product of their Laplace transforms we get (with usual notations) $Ee^{-tS_n} = (\int\limits_{[0,\infty)} e^{-tx} d\mu(x))^n = e^{-nct^\alpha} = Ee^{-t(n^{1/\alpha}X_1)}$ $\forall t$ which implies $S_n \stackrel{d}{=} n^{1/\alpha}X_1$. It follows that e^{-ct^α} is the Laplace transform of a positive (strictly) stable random variable.

Proposition 34

Let X be positive, strictly stable with index α and non-degenerate. Then $Ee^{-tX} = e^{-ct^\alpha}$ $(t > 0)$ for some $c > 0$.

Proof: let $f(t) = \log Ee^{-tX}$. Since $Ee^{-tn^{1/\alpha}X} = Ee^{-tS_n} = (Ee^{-tX})^n$ we get $f(n^{1/\alpha}t) = nf(t)$ $\forall n \geq 1$ $\forall t > 0$. Also f is continuous. It follows that $f(\frac{n^{1/\alpha}}{m^{1/\alpha}}t) = \frac{n}{m}f(t)$ $\forall n, m \geq 1$ and hence $f(s^{1/\alpha}t) = sf(t)$ $\forall s, t > 0$. Put $t = 1$ and replace s by s^α to get $f(s) = s^\alpha c$ where $c = f(1)$. Thus $Ee^{-tX} = e^{f(t)} = e^{ct^\alpha}$. Necessarily $c > 0$.

We are now ready to prove the proposition. Let μ be a stable measure with $\mu([0, \infty)) = 1$. [We already know that μ is absolutely continuous so $\mu\{0\} = 0$]. We know that $\alpha \in (0, 1)$. Let $\mu * \delta_{c_0}$ be strictly stable. Let $Y_n = X_n - c_0$. Then $Y_1 + Y_2 + \ldots + Y_n \stackrel{d}{=} n^{1/\alpha}Y_1$ and $Y_1 \geq -c_0$ so $Y_1 + Y_2 + \ldots + Y_n \geq -nc_0$ and

$n^{1/\alpha}Y_1 \geq -nc_0$ or $Y_1 \geq -n^{1-\frac{1}{\alpha}}c_0$. Letting $n \to \infty$ we get $Y_1 \geq 0$. Since Y_1 is non-negative, strictly stable and non-degenerate we have $Ee^{-tY_1} = e^{-ct^\alpha}$ for some $c > 0$. Hence $Ee^{-tX_1} = e^{-ct^\alpha}e^{tc_0}$ and the proposition is proved. [Since $Ee^{-tX_1} < 1$ we get $-ct^\alpha + tc_0 < 0$ $\forall t > 0$ which implies $c_0 \leq 0$].

Theorem 35

Let μ be stable with characteristic function with $\log\phi(t) = -c\,|t|^{\alpha}\,\{\cos\frac{\pi\alpha}{2} + i\beta s(t)\} + idt$ and $0<\alpha<1$. Then $\mu(0,\infty)=1$ if and only if $0<\alpha<1, b\geq 0$ and $\beta=1$.

Proof: suppose $\mu((0,\infty)) = 1$. Then $0<\alpha<1$. We have to show that $\beta=1$ and $b\geq 0$. We claim that the Levy measure ν is also concentrated on $(0,\infty)$. Suppose x_0 is in the support of ν so that $\nu(x_0-\varepsilon, x_0+\varepsilon) > 0\ \forall\varepsilon>0$. For any finite measure λ we write $e(\lambda)$ for the probability measure $e^{-\lambda(\mathbb{R})}\sum_{k=0}^{\infty}\frac{\lambda^{*(n)}}{n!}$ ($\lambda^{*(n)}$ being the $n-$ fold convolution of λ with itself). The characteristic function of $e(\lambda)$ at t is $e^{-\lambda(\mathbb{R})}\sum_{k=0}^{\infty}\int e^{itx}d\frac{\lambda^{*(n)}}{n!} = e^{-\lambda(\mathbb{R})}\sum_{k=0}^{\infty}\frac{1}{n!}(\int e^{itx}\lambda(x))^n = e^{-\lambda(\mathbb{R})}e^{\int e^{itx}\lambda(x)} = e^{\int\{e^{itx}-1\}\lambda(x)}$. Note that nx_0 belongs to the support of $e(\nu_1)$ where ν_1 is the restriction of ν to $\{x:\left|x>\frac{|x_0|}{2}\right|\}$. [This is because $\nu_1^{*(n)}(nx_0-\varepsilon, nx_0+\varepsilon)\geq\{\nu_1((x_0-\frac{\varepsilon}{n}, x_0+\frac{\varepsilon}{n})\}^n$]. Recall that $\phi(t) = e^{ict-\int(e^{itx}-1-\frac{itx}{1+x^2})d\nu(x)}$ for some constant c. Let μ_0 be the measure with characteristic function $e^{ic't-\int_{\{x:|x|\leq\frac{|x_0|}{2}\}}(e^{itx}-1-\frac{itx}{1+x^2})d\nu(x)}$ where $c' = c + \int_{\{x:|x|>\frac{|x_0|}{2}\}}\frac{x}{1+x^2}d\nu(x)$. Then $\mu=\mu_0 * e(\nu_1)$. [The characteristic functions of the two sides coincide]. If y belongs to the support of μ_0 then $y+nx_0$ belongs to the support of μ and hence $y+nx_0\geq 0\ \forall n$. Hence $x_0\geq 0$ proving that ν is supported by $(0,\infty)$. Since $d\nu(x) = C_1x^{-1-\alpha}I_{\{x>0\}}dx + C_2x^{-1-\alpha}I_{\{x<0\}}dx$ we get $C_2=0$ and $\beta=\frac{C_1-C_2}{C_1+C_2}=1$. Also $S_n - nb \overset{d}{=} n^{1/\alpha}(X_1-b)\geq -n^{1/\alpha}b$ so $S_n\geq nb-n^{1/\alpha}b$. This implies $X_1\geq(1-n^{1/\alpha-1})b$. If $b<0$ then $(1-n^{1/\alpha-1})b\to\infty$ which leads to the contradiction $X_1=\infty$ a.s..

Now we prove the converse. Suppose $0<\alpha<1, \beta=1$ and $b\geq 0$. We can write μ as $\mu_1 * \delta_b$ with μ_1 strictly stable. The Levy measure of μ_1 is the same as the Levy measure ν of μ. Since $\beta=1$ we get $C_2=0$ so $\nu(-\infty,0)=0$. The characteristic function ϕ_1 of μ_1 is given by $\phi_1(t) = e^{ib_1t+C_1\int_0^{\infty}(e^{itx}-1-\frac{itx}{1+x^2})x^{-1-\alpha}dx}$. Since μ_1 is strictly stable we get $\phi_1(t) = \lim e^{ib_1t+C_1\int_{1/n}^{\infty}(e^{itx}-1)x^{-1-\alpha}dx}$. If λ_n is the restriction of ν to $(\frac{1}{n},\infty)$ then $e(\lambda_n)(-\infty,0)=0$ and μ_1 is the weak limit of λ_n. Hence $\mu_1(-\infty,0)=0$. It follows that $\mu(-\infty,0)=\mu_1(-\infty,-b)=0$ because

$b \geq 0$.

Theorem 36

Let X and Y be independent strictly stable random variables with indices α and β respectively. Suppose $Y > 0$ a.s.. Then $XY^{1/\alpha}$ is $S(\alpha\beta)$.

Proof: if $\{X_n\}$ is i.i.d. with the same distribution as X and $\{Y_n\}$ is i.i.d., independent of $\{X_n\}$, with the same distribution as Y then $\{X_n Y_n^{1/\alpha}\}$ is i.i.d. with the same distribution as $XY^{1/\alpha}$. It suffices to show that $\sum_{j=1}^{n} X_j Y_j^{1/\alpha} \stackrel{d}{=} n^{1/(\alpha\beta)} XY^{1/\alpha}$ for each n. Recall that $aX_1 + bX_2 \stackrel{d}{=} (a^\alpha + b^\alpha)^{1/\alpha} X_1$. A simple induction argument shows that $\sum_{j=1}^{n} a_j X_j \stackrel{d}{=} (\sum_{j=1}^{n} a_j^\alpha)^{1/\alpha} X_1$ for any n and any positive numbers $a_1, a_2, \ldots, a_n$. Hence, the conditional distribution of $\sum_{j=1}^{n} X_j Y_{\mathrm{j}n}^{1/\alpha}$ given $Y_1, Y_2, \ldots, Y_n$ is equal to that of $(\sum_{j=1}^{n} Y_j)^{1/\alpha} X_1$ which is equal to the distribution of $(n^{1/\beta})^{1/\alpha} X_1 Y_1^{1/\alpha}$. The proof is completed by taking expectations.

Remark: an interesting special case is when $\alpha = 2$. If X and Y are i.i.d. with $N(0,1)$ distribution then $\frac{Y}{|X|}$ has Cauchy distribution.

Remark: stable have densities but these do not have a closed form. Series representations of the densities are available.

Remark: Marcus (Z.W., V 64, 1983, 139-156) showed that if $\phi(t,s) = e^{iar\cos(3\theta)} e^{-r^\alpha}$ where $0 < \alpha < 1$ and $t + is = re^{i\theta} (r > 0, \theta \in \mathbb{R})$ then ϕ is the characteristic function of a distribution on $\mathbb{R}^2$ which is *not* stable but its marginals both have stable distribution with index α.

Series Representation of stable laws [due to Le Page]

Lemma 37

Let $T_j, 1 \leq j \leq N+1$ be i.i.d. exponential with parameter 1. Let $S = T_1 + T_2 + \ldots + T_{N+1}$. Then $(\frac{T_1}{S}, \frac{T_2}{S}, \ldots, \frac{T_N}{S})$ has an absolutely continuous distribution whose density f is given by $f(x_1, x_2, \ldots, x_N) = n! I_{[0,\infty)}(x_1) I_{[0,\infty)}(x_2) \ldots I_{[0,\infty)}(x_N) I_{\{x_1+x_2+\ldots+x_N \leq 1\}}$.

We leave the proof as an exercise. [Write down the joint distribution of $\frac{T_1}{S}, \frac{T_2}{S}, \ldots, \frac{T_N}{S}, S$, make a suitable change of variable and use the formula

$$\det \begin{pmatrix} x_{N+1} & 0 & . & . & x_1 \\ 0 & x_{N+1} & . & . & x_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ -x_{N+1} & -x_{N+1} & . & . & 1-\sum_{j=1}^{N} x_j \end{pmatrix} = x_{N+1}^N \text{ a seen by the first } N$$

rows to the last row.

Theorem 38

Let $\Gamma_1, \Gamma_2, \ldots$ be the arrival times in a Poisson process with parameter 1. Let $\{\varepsilon_j\}$ be symmetric $\{-1, 1\}$ valued random variables independent of $\Gamma_j' s$. Let $0 < \alpha < 2$. Then $\sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$ converges a.s. to an $S(\alpha)$ random variable.

Proof of the theorem: $P\{\sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$ converges $/\Gamma_1, \Gamma_2,,,\} = I_{\{\sum_{j=1}^{\infty} \Gamma^{-2/\alpha} < \infty\}}$ by the three-series-theorem (or by the fact that $\mathbb{R}$ is of type 2 and cotype 2). By Strong Law of Large Numbers $\frac{\Gamma_j}{j} \to 1$ a.s.. Hence $\sum_{j=1}^{\infty} \Gamma^{-2/\alpha} < \infty$ a.s. and $\sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$ converges a.s.. Let $\{U_n\}$ be i.i.d., independent of $\{\varepsilon_j\}$, having uniform distribution on $(0, 1)$. Let $X_j = \varepsilon_j U_j^{-1/\alpha}$. Then $\{X_j\}$ is i.i.d. and we claim that $N^{-1/\alpha} \sum_{j=1}^{N} X_j \stackrel{d}{=} (\frac{\Gamma_{N+1}}{N})^{1/\alpha} \sum_{j=1}^{N} \varepsilon_j \Gamma_j^{-1/\alpha}$. Once this is proved the theorem follows easily: $\frac{1}{(Nk)^{1/\alpha}} \sum_{j=1}^{Nk} X_j = \frac{1}{k^{1/\alpha}} \{\frac{1}{N^{1/\alpha}} \sum_{j=1}^{N} X_j\} + \frac{1}{k^{1/\alpha}} \{\frac{1}{N^{1/\alpha}} \sum_{j=N+1}^{2N} X_j\} + \ldots + \frac{1}{k^{1/\alpha}} \{\frac{1}{N^{1/\alpha}} \sum_{j=N(k-1)+1}^{Nk} X_j\}$; since $(\frac{\Gamma_{N+1}}{N})^{1/\alpha} \sum_{j=1}^{N} \varepsilon_j \Gamma_j^{-1/\alpha} \xrightarrow{d} \sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$ as $N \to \infty$ we get $R \stackrel{d}{=} \frac{1}{k^{1/\alpha}} R_1 + \frac{1}{k^{1/\alpha}} R_2 + \ldots + \frac{1}{k^{1/\alpha}} R_k$ where $\{R_j\}$ is i.i.d. with same distribution as $\sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$. This proves that the symmetric random variable $\sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$ has stable distribution with parameter α. To complete the proof we now show that $N^{-1/\alpha} \sum_{j=1}^{N} \varepsilon_j U_j^{-1/\alpha} \stackrel{d}{=} (\frac{\Gamma_{N+1}}{N})^{1/\alpha} \sum_{j=1}^{N} \varepsilon_j \Gamma_j^{-1/\alpha}$. We begin with the following fact: if $\{U_{(1)}, U_{(2)}, \ldots, U_{(N)})$ is the order statistics from

$\{U_1, U_2, \ldots, U_n\}$ then $(U_{(1)}, U_{(2)}, \ldots, U_{(N)})) \overset{d}{=} (\frac{\Gamma_1}{\Gamma_{N+1}}, \frac{\Gamma_2}{\Gamma_{N+1}}, \ldots, \frac{\Gamma_N}{\Gamma_{N+1}})$. This is easy: write down the joint distribution of $\frac{T_1}{S}, \frac{T_2}{S}, \ldots, \frac{T_N}{S}, S$ as above and make a change of variable. [Show that the joint density of $(\frac{\Gamma_1}{S}, \frac{\Gamma_2}{S}, \ldots, \frac{\Gamma_N}{S})$ is $N! I_{|0<x_1<x_2<\ldots<x_n<1\}}$ which is also the joint density of $(U_{(1)}, U_{(2)}, \ldots, U_{(N)}))]$.

Assuming that $\varepsilon_j' s, \Gamma_j' s$ and $U_j' s$ are independent of each other we get $N^{-1/\alpha} \sum_{j=1}^{N} X_j =$

$$N^{-1/\alpha} \sum_{j=1}^{N} \varepsilon_j U_j^{-1/\alpha} \overset{d}{=} N^{-1/\alpha} \sum_{j=1}^{N} \varepsilon_j U_{(j)}^{-1/\alpha} \overset{d}{=} N^{-1/\alpha} \sum_{j=1}^{N} \varepsilon_j (\frac{\Gamma_j}{\Gamma_{N+1}})^{-1/\alpha}$$

$$= N^{-1/\alpha} \sum_{j=1}^{N} \varepsilon_j \Gamma_j^{-1/\alpha} \Gamma_{N+1}^{1/\alpha} = \sum_{j=1}^{N} \varepsilon_j \Gamma_j^{-1/\alpha} (\frac{\Gamma_{N+1}}{N})^{1/\alpha}$$ which finishes the proof.

Corollary 39

Let X have a (non-degenerate) symmetric stable distribution with index α. Then $X \overset{d}{=} c \sum_{j=1}^{\infty} \varepsilon_j \Gamma_j^{-1/\alpha}$ for some $c > 0$

Proof of corollary: $E e^{itX} = e^{-c|t|^\alpha}$ for some $c > 0$ and hence the corollary is immediate.

Exercise

Using arguments similar to the above show that $\sum_{j=1}^{\infty} \Gamma_j^{-1/\alpha}$ converges a.s. and has a positive stable distribution with parameter α.