Lectures on Lévy Processes and Stochastic Calculus, Braunschweig, Lecture 1: Infinite Divisibility

David Applebaum

Probability and Statistics Department, University of Sheffield, UK

July 22nd - 24th 2010

Introduction

A **Lévy process** is essentially a stochastic process with stationary and independent increments.

The basic theory was developed, principally by Paul Lévy in the 1930s. In the past 20 years there has been a renaissance of interest and a plethora of books, articles and conferences. Why ?

There are both theoretical and practical reasons.

Theoretical

- Lévy processes are simplest generic class of process which have (a.s.) continuous paths interspersed with random jumps of arbitrary size occurring at random times.
- Lévy processes comprise a natural subclass of *semimartingales* and of *Markov processes of Feller type*.
- There are many interesting examples - Brownian motion, simple and compound Poisson processes, α-stable processes, subordinated processes, financial processes, relativistic process, Riemann zeta process . . .

- Noise. Lévy processes are a good model of “noise” in random dynamical systems.

  Input + Noise = Output

  Attempts to describe this differentially leads to *stochastic calculus*. A large class of Markov processes can be built as solutions of *stochastic differential equations* (SDEs) driven by Lévy noise.

  Lévy driven *stochastic partial differential equations* (SPDEs) are currently being studied with some intensity.
- Robust structure. Most applications utilise Lévy processes taking values in Euclidean space but this can be replaced by a Hilbert space, a Banach space (both of these are important for SPDEs), a locally compact group, a manifold. Quantised versions are non-commutative Lévy processes on quantum groups.

Applications

These include:

- Turbulence via Burger's equation (Bertoin)
- New examples of quantum field theories (Albeverio, Gottshalk, Wu)
- Viscoelasticity (Bouleau)
- Time series - Lévy driven CARMA models (Brockwell, Marquardt)
- Stochastic resonance in non-linear signal processing (Patel, Kosco, Applebaum)
- Finance and insurance (a cast of thousands)

The biggest explosion of activity has been in mathematical finance. Two major areas of activity are:

- option pricing in incomplete markets.
- interest rate modelling.

Some Basic Ideas of Probability

Notation. Our state space is Euclidean space $\mathbb{R}^d$. The inner product between two vectors $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ is

$$(x, y) = \sum_{i=1}^{d} x_i y_i.$$

The associated norm (length of a vector) is

$$|x| = (x, x)^{\frac{1}{2}} = \left(\sum_{i=1}^{d} x_i^2\right)^{\frac{1}{2}}.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space, so that Ω is a set, $\mathcal{F}$ is a σ-algebra of subsets of Ω and P is a probability measure defined on $(\Omega, \mathcal{F})$.

Random variables are measurable functions $X : \Omega \rightarrow \mathbb{R}^d$. The law of X is p_X, where for each $A \in \mathcal{F}, p_X(A) = P(X \in A)$.

$(X_n, n \in \mathbb{N})$ are *independent* if for all $i_1, i_2, \ldots i_r \in \mathbb{N}, A_{i_1}, A_{i_2}, \ldots, A_{i_r} \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned} & P(X_{i_1} \in A_1, X_{i_2} \in A_2, \ldots, X_{i_r} \in A_r) \\ = \; & P(X_{i_1} \in A_1) P(X_{i_2} \in A_2) \cdots P(X_{i_r} \in A_r). \end{aligned}$$

If X and Y are independent, the law of $X + Y$ is given by *convolution of measures*

$$p_{X+Y} = p_X * p_Y, \text{ where } (p_X * p_Y)(A) = \int_{\mathbb{R}^d} p_X(A - y) p_Y(dy).$$

Equivalently

$$\int_{\mathbb{R}^d} g(y)(p_X * p_Y)(dy) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(x + y) p_X(dx) p_Y(dy),$$

for all $g \in B_b(\mathbb{R}^d)$ (the bounded Borel measurable functions on $\mathbb{R}^d$).
If X and Y are independent with densities f_X and f_Y, respectively, then $X + Y$ has density f_{X+Y} given by *convolution of functions*:

$$f_{X+Y} = f_X * f_Y, \text{ where } (f_X * f_Y)(x) = \int_{\mathbb{R}^d} f_X(x - y) f_Y(y) dy.$$

The *characteristic function* of X is $\phi_X : \mathbb{R}^d \to \mathbb{C}$, where

$$\phi_X(u) = \int_{\mathbb{R}^d} e^{i(u,x)} p_X(dx).$$

Theorem (Kac's theorem)

$X_1, \ldots, X_n$ are independent if and only if

$$\mathbb{E}\left(\exp\left(i\sum_{j=1}^{n}(u_j, X_j)\right)\right) = \phi_{X_1}(u_1)\cdots\phi_{X_n}(u_n)$$

for all $u_1, \ldots, u_n \in \mathbb{R}^d$.

More generally, the characteristic function of a probability measure μ on $\mathbb{R}^d$ is

$$\phi_\mu(u) = \int_{\mathbb{R}^d} e^{i(u,x)} \mu(dx).$$

Important properties are:-

1. $\phi_\mu(0) = 1$.
2. ϕ_μ is *positive definite* i.e. $\sum_{i,j} c_i \bar{c}_j \phi_\mu(u_i - u_j) \geq 0$, for all $c_i \in \mathbb{C}, u_i \in \mathbb{R}^d, 1 \leq i, j \leq n, n \in \mathbb{N}$.
3. ϕ_μ is uniformly continuous - Hint: Look at $|\phi_\mu(u+h) - \phi_\mu(u)|$ and use dominated convergence)).

Also $\mu \to \phi_\mu$ is injective.

Conversely *Bochner's theorem* states that if $\phi : \mathbb{R}^d \to \mathbb{C}$ satisfies (1), (2) and is continuous at $u = 0$, then it is the characteristic function of some probability measure μ on $\mathbb{R}^d$.

$\psi : \mathbb{R}^d \to \mathbb{C}$ is *conditionally positive definite* if for all $n \in \mathbb{N}$ and $c_1, \ldots, c_n \in \mathbb{C}$ for which $\sum_{j=1}^n c_j = 0$ we have

$$\sum_{j,k=1}^{n} c_j \bar{c}_k \psi(u_j - u_k) \geq 0,$$

for all $u_1, \ldots, u_n \in \mathbb{R}^d$.

Note: *conditionally positive definite* is sometimes called *negative definite*.

$\psi : \mathbb{R}^d \to \mathbb{C}$ will be said to be *hermitian* if $\overline{\psi(u)} = \psi(-u)$, for all $u \in \mathbb{R}^d$.

Theorem (Schoenberg correspondence)

$\psi : \mathbb{R}^d \to \mathbb{C}$ is hermitian and conditionally positive definite if and only if $e^{t\psi}$ is positive definite for each $t > 0$.

Proof. We only give the easy part here.
Suppose that $e^{t\psi}$ is positive definite for all $t > 0$. Fix $n \in \mathbb{N}$ and choose $c_1, \ldots, c_n$ and $u_1, \ldots, u_n$ as above.
We then find that for each $t > 0$,

$$\frac{1}{t}\sum_{j,k=1}^{n} c_j \bar{c}_k (e^{t\psi(u_j - u_k)} - 1) \geq 0,$$

and so

$$\sum_{j,k=1}^{n} c_j \bar{c}_k \psi(u_j - u_k) = \lim_{t \to 0} \frac{1}{t} \sum_{j,k=1}^{n} c_j \bar{c}_k (e^{t\psi(u_j - u_k)} - 1) \geq 0.$$

□

Infinite Divisibility

We study this first because a Lévy process is infinite divisibility in motion, i.e. infinite divisibility is the underlying probabilistic idea which a Lévy process embodies dynamically.

Let μ be a probability measure on $\mathbb{R}^d$. Define $\mu^{*^n} = \mu * \cdots * \mu$ (n times). We say that μ has a *convolution nth root*, if there exists a probability measure $\mu^{\frac{1}{n}}$ for which $(\mu^{\frac{1}{n}})^{*^n} = \mu$.

μ is *infinitely divisible* if it has a convolution nth root for all $n \in \mathbb{N}$. In this case $\mu^{\frac{1}{n}}$ is unique.

Theorem

μ is infinitely divisible iff for all $n \in \mathbb{N}$, there exists a probability measure μ_n with characteristic function ϕ_n such that

$$\phi_\mu(u) = (\phi_n(u))^n,$$

for all $u \in \mathbb{R}^d$. Moreover $\mu_n = \mu^{\frac{1}{n}}$.

Proof. If μ is infinitely divisible, take $\phi_n = \phi_{\mu^{\frac{1}{n}}}$. Conversely, for each $n \in \mathbb{N}$, by Fubini's theorem,

$$\begin{aligned} \phi_\mu(u) &= \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} e^{i(u, y_1 + \cdots + y_n)} \mu_n(dy_1) \cdots \mu_n(dy_n) \\ &= \int_{\mathbb{R}^d} e^{i(u,y)} \mu_n^{*^n}(dy). \end{aligned}$$

But $\phi_\mu(u) = \int_{\mathbb{R}^d} e^{i(u,y)} \mu(dy)$ and ϕ determines μ uniquely. Hence $\mu = \mu_n^{*^n}$. □

- If μ and ν are each infinitely divisible, then so is $\mu * \nu$.

- If $(\mu_n, n \in \mathbb{N})$ are infinitely divisible and $\mu_n \stackrel{w}{\Rightarrow} \mu$, then μ is infinitely divisible.

[Note: *Weak convergence.* $\mu_n \stackrel{w}{\Rightarrow} \mu$ means

$$\lim_{n \to \infty} \int_{\mathbb{R}^d} f(x) \mu_n(dx) = \int_{\mathbb{R}^d} f(x) \mu(dx),$$

for each $f \in C_b(\mathbb{R}^d)$.]

A random variable X is *infinitely divisible* if its law p_X is infinitely divisible, $X \stackrel{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, where $Y_1^{(n)}, \ldots, Y_n^{(n)}$ are i.i.d., for each $n \in \mathbb{N}$.

Examples of Infinite Divisibility

In the following, we will demonstrate infinite divisibility of a random variable X by finding i.i.d. $Y_1^{(n)}, \ldots, Y_n^{(n)}$ such that $X \stackrel{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, for each $n \in \mathbb{N}$.

Example 1 - Gaussian Random Variables

Let $X = (X_1, \ldots, X_d)$ be a random vector.
We say that it is $(non-degenerate) Gaussian$ if it there exists a vector $m \in \mathbb{R}^d$ and a strictly positive-definite symmetric $d \times d$ matrix A such that X has a pdf (probability density function) of the form:-

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det(A)}} \exp\left(-\frac{1}{2}(x - m, A^{-1}(x - m))\right), \qquad (1.1)$$

for all $x \in \mathbb{R}^d$.
In this case we will write $X \sim N(m, A)$. The vector m is the mean of X, so $m = \mathbb{E}(X)$ and A is the covariance matrix so that $A = \mathbb{E}((X - m)(X - m)^T)$.

A standard calculation yields

$$\phi_X(u) = \exp\{i(m,u) - \frac{1}{2}(u, Au)\}, \quad (1.2)$$

and hence

$$(\phi_X(u))^{\frac{1}{n}} = \exp\left\{i\left(\frac{m}{n}, u\right) - \frac{1}{2}\left(u, \frac{1}{n}Au\right)\right\},$$

so we see that X is infinitely divisible with each $Y_j^{(n)} \sim N(\frac{m}{n}, \frac{1}{n}A)$ for each $1 \leq j \leq n$.
We say that X is a *standard normal* whenever $X \sim N(0, \sigma^2 I)$ for some $\sigma > 0$.

We say that X is *degenerate Gaussian* if (1.2) holds with $\det(A) = 0$, and these random variables are also infinitely divisible.

Example 2 - Poisson Random Variables

In this case, we take $d = 1$ and consider a random variable X taking values in the set $n \in \mathbb{N} \cup \{0\}$. We say that is *Poisson* if there exists $c > 0$ for which

$$P(X = n) = \frac{c^n}{n!}e^{-c}.$$

In this case we will write $X \sim \pi(c)$. We have $\mathbb{E}(X) = \text{Var}(X) = c$. It is easy to verify that

$$\phi_X(u) = \exp[c(e^{iu} - 1)],$$

from which we deduce that X is infinitely divisible with each $Y_j^{(n)} \sim \pi(\frac{c}{n})$, for $1 \leq j \leq n, n \in \mathbb{N}$.

Example 3 - Compound Poisson Random Variables

Let $(Z(n), n \in \mathbb{N})$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}^d$ with common law μ_Z and let $N \sim \pi(c)$ be a Poisson random variable which is independent of all the $Z(n)$'s.The *compound Poisson random variable* X is defined as follows:-

$$X := \begin{cases} 0 & \text{if } N = 0 \\ Z(1) + \cdots + Z(N) & \text{if } N > 0. \end{cases}$$

Theorem

For each $u \in \mathbb{R}^d$,

$$\phi_X(u) = \exp\left[\int_{\mathbb{R}^d}(e^{i(u,y)} - 1)c\mu_Z(dy)\right].$$

Proof. Let ϕ_Z be the common characteristic function of the Z_n's. By conditioning and using independence we find,

$$\begin{aligned}\phi_X(u) &= \sum_{n=0}^{\infty}\mathbb{E}(e^{i(u,Z(1)+\cdots+Z(N))}|N = n)P(N = n) \\ &= \sum_{n=0}^{\infty}\mathbb{E}(e^{i(u,Z(1))+\cdots+Z(n))})e^{-c}\frac{c^n}{n!} \\ &= e^{-c}\sum_{n=0}^{\infty}\frac{[c\phi_Z(u)]^n}{n!} \\ &= \exp[c(\phi_Z(u) - 1)],\end{aligned}$$

and the result follows on writing $\phi_Z(u) = \int e^{i(u,y)}\mu_Z(dy)$. □

If X is compound Poisson as above, we write $X \sim \pi(c, \mu_Z)$. It is clearly infinitely divisible with each $Y_j^{(n)} \sim \pi(\frac{c}{n}, \mu_Z)$, for $1 \leq j \leq n$.

The Lévy-Khintchine Formula

de Finetti (1920's) suggested that the most general infinitely divisible random variable could be written $X = Y + W$, where Y and W are independent, $Y \sim N(m, A)$, $W \sim \pi(c, \mu_Z)$. Then $\phi_X(u) = e^{\eta(u)}$, where

$$\eta(u) = i(m,u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d} (e^{i(u,y)} - 1) c\mu_Z(dy). \quad (1.3)$$

This is WRONG! $\nu(\cdot) = c\mu_Z(\cdot)$ is a finite measure here.
Lévy and Khintchine showed that ν can be σ-finite, provided it is what is now called a *Lévy measure* on $\mathbb{R}^d - \{0\} = \{x \in \mathbb{R}^d, x \neq 0\}$, i.e.

$$\int (|y|^2 \wedge 1)\nu(dy) < \infty, \quad (1.4)$$

(where $a \wedge b := \min\{a, b\}$, for $a, b \in \mathbb{R}$). Since $|y|^2 \wedge \epsilon \leq |y|^2 \wedge 1$ whenever $0 < \epsilon \leq 1$, it follows from (1.4) that

$$\nu((-\epsilon, \epsilon)^c) < \infty \quad \text{for all } \epsilon > 0.$$

Here is the fundamental result of this lecture:-

Theorem (Lévy-Khintchine)

A Borel probability measure μ on $\mathbb{R}^d$ is infinitely divisible if there exists a vector $b \in \mathbb{R}^d$, a non-negative symmetric $d \times d$ matrix A and a Lévy measure ν on $\mathbb{R}^d - \{0\}$ such that for all $u \in \mathbb{R}^d$,

$$\phi_\mu(u) = \exp\left[i(b,u) - \frac{1}{2}(u, Au)\right. \quad (1.5)$$

$$\left. + \int_{\mathbb{R}^d - \{0\}} (e^{i(u,y)} - 1 - i(u,y)\mathbf{1}_{\hat{B}}(y))\nu(dy)\right]. \quad (1.6)$$

where $\hat{B} = B_1(0) = \{y \in \mathbb{R}^d; |y| < 1\}$.
Conversely, any mapping of the form (1.5) is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}^d$.

The triple (b, A, ν) is called the *characteristics* of the infinitely divisible random variable X. Define $\eta = \log \phi_\mu$, where we take the principal part of the logarithm. η is called the *Lévy symbol*.

η may also be called the *characteristic exponent*.
We're not going to prove this result here. To understand it, it is instructive to let $(U_n, n \in \mathbb{N})$ be a sequence of Borel sets in $B_1(0)$ with $U_n \downarrow \{0\}$. Observe that

$$\eta(u) = \lim_{n \to \infty} \eta_n(u) \quad \text{where each}$$

$$\eta_n(u) = i\left[\left(b - \int_{U_n^c \cap \hat{B}} y\nu(dy), u\right)\right] - \frac{1}{2}(u, Au) + \int_{U_n^c} (e^{i(u,y)} - 1)\nu(dy),$$

so η is in some sense (to be made more precise later) the limit of a sequence of sums of Gaussians and independent compound Poissons. Interesting phenomena appear in the limit as we'll see below.

First, we classify Lévy symbols analytically:-

Theorem

η is a Lévy symbol if and only if it is a continuous, hermitian conditionally positive definite function for which $\eta(0) = 0$.

Hilbert Space. The Lévy-Khintchine formula extends in an obvious manner to separable Hilbert spaces H. We interpret $(\cdot, \cdot)$ as the Hilbert space inner product. The characteristics (b, A, ν) are defined similarly - A is a positive, symmetric linear operator on H.

Banach space. We can also extend to a separable Banach space E having dual E^*. Define $\Phi_\mu : E^* \to \mathbb{C}$ by

$$\Phi_\mu(u) = \int_E e^{i\langle u,x\rangle}\mu(dx),$$

where $\langle \cdot, \cdot \rangle$ is duality. Replacing $(\cdot, \cdot)$ with $\langle \cdot, \cdot \rangle$ the Lévy-Khintchine formula extends naturally but care must be taken in defining a Lévy measure!

Stable Laws

This is one of the most important subclasses of infinitely divisible laws. We consider the general central limit problem in dimension $d = 1$, so let $(Y_n, n \in \mathbb{N})$ be a sequence of real valued i.i.d. random variables and consider the rescaled partial sums

$$S_n = \frac{Y_1 + Y_2 + \cdots + Y_n - b_n}{\sigma_n},$$

where $(b_n, n \in \mathbb{N})$ is an arbitrary sequence of real numbers and $(\sigma_n, n \in \mathbb{N})$ an arbitrary sequence of positive numbers. We are interested in the case where there exists a random variable X for which

$$\lim_{n \to \infty} P(S_n \leq x) = P(X \leq x), \qquad (1.7)$$

for all $x \in \mathbb{R}$, i.e. $(S_n, n \in \mathbb{N})$ converges in distribution to X. If each $b_n = nm$ and $\sigma_n = \sqrt{n}\sigma$ for fixed $m \in \mathbb{R}, \sigma > 0$ then $X \sim N(m, \sigma^2)$ by the usual Laplace - de-Moivre central limit theorem.

More generally a random variable is said to be *stable* if it arises as a limit as in (1.7). It is not difficult to show that (1.7) is equivalent to the following:-

There exist real valued sequences $(c_n, n \in \mathbb{N})$ and $(d_n, n \in \mathbb{N})$ with each $c_n > 0$ such that

$$X_1 + X_2 + \cdots + X_n \stackrel{d}{=} c_n X + d_n \qquad (1.8)$$

where $X_1, \ldots, X_n$ are independent copies of X. X is said to be *strictly stable* if each $d_n = 0$.
To see that (1.8) $\Rightarrow$ (1.7) take each $Y_j = X_j, b_n = d_n$ and $\sigma_n = c_n$. In fact it can be shown that the only possible choice of c_n in (1.8) is $c_n = \sigma n^{\frac{1}{\alpha}}$, where $0 < \alpha \leq 2$ and $\sigma > 0$. The parameter α plays a key role in the investigation of stable random variables and is called the *index of stability*.
Note that (1.8) can also be expressed in the equivalent form

$$\phi_X(u)^n = e^{iud_n}\phi_X(c_n u),$$

for each $u \in \mathbb{R}$.

It follows immediately from (1.8) that all stable random variables are infinitely divisible and the characteristics in the Lévy-Khintchine formula are given as follows:

Theorem

If X is a stable real-valued random variable, then its characteristics must take one of the two following forms.

1. *When $\alpha = 2$, $\nu = 0$ (so $X \sim N(b, A)$).*
2. *When $\alpha \neq 2$, $A = 0$ and*
   $\nu(dx) = \frac{c_1}{x^{1+\alpha}} \mathbf{1}_{(0,\infty)}(x)dx + \frac{c_2}{|x|^{1+\alpha}} \mathbf{1}_{(-\infty,0)}(x)dx,$
   where $c_1 \geq 0, c_2 \geq 0$ and $c_1 + c_2 > 0$.

A careful transformation of the integrals in the Lévy-Khintchine formula gives a different form for the characteristic function which is often more convenient.

Theorem

A real-valued random variable X is stable if and only if there exists $\sigma > 0, -1 \leq \beta \leq 1$ and $\mu \in \mathbb{R}$ such that for all $u \in \mathbb{R}$,

1. $$\phi_X(u) = \exp\left[i\mu u - \frac{1}{2}\sigma^2 u^2\right]$$
   when $\alpha = 2$
2. $$\phi_X(u) = \exp\left[i\mu u - \sigma^\alpha |u|^\alpha \left(1 - i\beta sgn(u) \tan(\frac{\pi\alpha}{2})\right)\right]$$
   when $\alpha \neq 1, 2$.

Theorem

3
$$\phi_X(u) = \exp\left[i\mu u - \sigma|u|\left(1 + i\beta\frac{2}{\pi}sgn(u)\log(|u|)\right)\right] \quad \textit{when } \alpha = 1.$$

It can be shown that $\mathbb{E}(X^2) < \infty$ if and only if $\alpha = 2$ (i.e. X is Gaussian) and $\mathbb{E}(|X|) < \infty$ if and only if $1 < \alpha \leq 2$.
All stable random variables have densities f_X, which can in general be expressed in series form. In three important cases, there are closed forms.

1. **The Normal Distribution**
   $$\alpha = 2, \quad X \sim N(\mu, \sigma^2).$$
2. **The Cauchy Distribution**
   $$\alpha = 1, \beta = 0 \quad f_X(x) = \frac{\sigma}{\pi[(x-\mu)^2 + \sigma^2]}.$$
3. **The Lévy Distribution**
   $$\alpha = \frac{1}{2}, \beta = 1 \quad f_X(x) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{\frac{3}{2}}} \exp\left(-\frac{\sigma}{2(x-\mu)}\right),$$
   for $x > \mu$.

In general the series representations are given in terms of a real valued parameter λ.

For $x > 0$ and $0 < \alpha < 1$:

$$f_X(x, \lambda) = \frac{1}{\pi x}\sum_{k=1}^{\infty}\frac{\Gamma(k\alpha + 1)}{k!}(-x^{-\alpha})^k \sin\left(\frac{k\pi}{2}(\lambda - \alpha)\right)$$

For $x > 0$ and $1 < \alpha < 2$,

$$f_X(x, \lambda) = \frac{1}{\pi x}\sum_{k=1}^{\infty}\frac{\Gamma(k\alpha^{-1} + 1)}{k!}(-x)^k \sin\left(\frac{k\pi}{2\alpha}(\lambda - \alpha)\right)$$

In each case the formula for negative x is obtained by using

$$f_X(-x, \lambda) = f_X(x, -\lambda), \quad \text{for } x > 0.$$

Note that if a stable random variable is symmetric then Theorem 8 yields

$$\phi_X(u) = \exp(-\rho^\alpha |u|^\alpha) \text{ for all } 0 < \alpha \leq 2, \quad (1.9)$$

where $\rho = \sigma(0 < \alpha < 2)$ and $\rho = \frac{\sigma}{\sqrt{2}}$, when $\alpha = 2$, and we will write $X \sim S\alpha S$ in this case.

One of the reasons why stable laws are so important in applications is the nice decay properties of the tails. The case $\alpha = 2$ is special in that we have exponential decay, indeed for a standard normal X there is the elementary estimate

$$P(X > y) \sim \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}y} \text{ as } y \to \infty,$$

When $\alpha \neq 2$ we have the slower polynomial decay as expressed in the following,

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

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

where $C_\alpha > 1$. The relatively slow decay of the tails for non-Gaussian stable laws ("heavy tails") makes them ideally suited for modelling a wide range of interesting phenomena, some of which exhibit "long-range dependence".

Deeper mathematical investigations of heavy tails require the mathematical technique of *regular variation*.

The generalisation of stability to random vectors is straightforward - just replace $X_1, \ldots, X_n$, X and each d_n in (1.8) by vectors and the formula in Theorem 7 extends directly. Note however that when $\alpha \neq 2$ in the random vector version of Theorem 7, the Lévy measure takes the form

$$\nu(dx) = \frac{c}{|x|^{d+\alpha}}dx$$

where $c > 0$.

We can generalise the definition of stable random variables if we weaken the conditions on the random variables $(Y(n), n \in \mathbb{N})$ in the general central limit problem by requiring these to be independent, but no longer necessarily identically distributed.
In this case (subject to a technical growth restriction) the limiting random variables are called *self-decomposable* (or of *class L*) and they are also infinitely divisible.
Alternatively a random variable X is self-decomposable if and only if for each $0 < a < 1$, there exists a random variable Y_a which is independent of X such that

$$X \stackrel{d}{=} aX + Y_a \Leftrightarrow \phi_X(u) = \phi_X(au)\phi_{Y_a}(u),$$

for all $u \in \mathbb{R}^d$.

An infinitely divisible law is self-decomposable if and only if the Lévy measure is of the form:

$$\nu(dx) = \frac{k(x)}{|x|}dx,$$

where k is decreasing on $(0, \infty)$ and increasing on $(-\infty, 0)$. There has recently been increasing interest in these distributions both from a theoretical and applied perspective. Examples include gamma, Pareto, Student-t, F and log-normal distributions.

Heavy Tails - Subexponentiality and Regular Variation

Consider the aggregate behaviour of the sum of a large number of i.i.d. random variables,

- This can be co-operative so that no one r.v. dominates $\rightarrow$ familiar Gaussian C.L.T. .
- This can be dominated by the behaviour of a single random variable $\rightarrow$ C.L.T. for stable laws.

To capture the second type of behaviour, we say that a non-negative random variable X is *subexponential* if as $x \rightarrow \infty$

$$P(X_1 + \cdots + X_n > x) \sim P(\max\{X_1, \ldots, X_n\} > x),$$

where $X_1, \ldots, X_n$ are independent copies of X.
A natural and more easily handled subclass of subexponential random variables are those of *regular variation of index* α.
$X \in \mathcal{R}_{-\alpha}$ where $\alpha > 0$ if

$$\lim_{x \to \infty} \frac{\overline{F_X}(cx)}{\overline{F_X}(x)} = c^{\alpha},$$

where $\overline{F_X}(x) := P(X > x)$.

e.g. X *Pareto* - parameters $K, \alpha > 0$, $\overline{F_X}(x) = \left(\frac{K}{K+x}\right)^{\alpha}$. X is self-decomposable.

Intuition: “Heavy tails” (subexponentiality) are due to “large jumps”.

Large jumps are governed by the tail of the Lévy measure (see Lecture 3)

Fact (Tail equivalence): If X is infinitely divisible then $\overline{F_X} \in \mathcal{R}_{-\alpha}$ if and only if $\overline{\nu} \in \mathcal{R}_{-\alpha}$.

In this case $\lim_{x \to \infty} \frac{\overline{F_X}(x)}{\overline{\nu}(x)} = 1$.