Summary : summing $X_1 + X_2 + \cdots X_n \equiv S_n$, with $p(x) \propto \frac{1}{|x|^{1+\mu}}$, $x \to \pm\infty$ converges, under appropriate shifting / rescaling, to some $L_{\beta,\mu}$, Lévy-stable
The rescaling is ≠t from CLT one :

$$\xi_n \equiv \frac{S_n - n\langle x \rangle}{n^{1/\mu}} \longrightarrow L_{\beta,\mu}(\xi) \quad , \quad 1 < \mu < 2$$

$$\xi_n \equiv \frac{S_n}{n^{1/\mu}} \longrightarrow \text{some other } L_{\beta,\mu}(\xi) \quad ; \quad 0 < \mu < 1$$

Cases $\mu = 1$ and $\mu = 2$ are special, see below for emergence of $\log n$ terms.
Rk : for $1 < \mu < 2$, a law of large numbers thus exists

$$\frac{S_n}{n} = \langle x \rangle + \underbrace{n^{1/\mu - 1}}_{\xrightarrow[n\to\infty]{} 0} \xi_n$$

but for $0 < \mu < 1$: $\frac{S_n}{n} = n^{1/\mu - 1} \xi_n$, explodes !
$S_n \propto n^{1/\mu}$, grows faster than linearly with n

4°b) Why all this ? Relation to extreme value statistics

Above results can be rationalized studying the statistics of $M_n = \max(X_1, \ldots, X_n)$

$$\begin{aligned} \Pr(M_n \leq m) &= \Pr(X_1 \leq m \text{ and } X_2 \leq m \text{ and} \cdots \text{and } X_n \leq m) \\ &= \left[\Pr(X_1 \leq m)\right]^n \\ &= \left[\int_0^m p(x)\,dx\right]^n \end{aligned}$$

This yields the pdf

$$p_{M_n}(m) = n\,p(m)\left[\int_0^m p(x)\,dx\right]^{n-1}$$

that we could maximize wrt to m to find the most probable / typical $M_n \to$ denoted M_n^*. We follow a more intuitive route.
M_n^* should not be too small nor too large : it is s.t

$\Pr(M_n \leq M_n^*)$ is of order 1, say $\frac{1}{2}$, or $\frac{1}{3}$, or 0,9 ... all the same

$$\Rightarrow \left[1 - \int_{M_n^*}^{\infty} p(x)\,dx\right]^n \xrightarrow[n\to\infty]{} \frac{1}{2}, \text{ say a cst}$$

$$\Rightarrow \int_{M_n^*}^{\infty} p(x)\,dx \equiv \Pr(M_n \geq M_n^*) \propto \frac{1}{n} \quad \text{for } n \to \infty$$

↳ very useful in practice

If $p(x) \underset{\infty}{\propto} \frac{1}{x^{1+\mu}}$, $\Pr(X \geq x) \propto \frac{1}{x^{\mu}}$

$$\Rightarrow \frac{1}{n} \propto \frac{1}{(M_n^*)^{\mu}} \quad \Rightarrow \quad \boxed{M_n^* \propto n^{1/\mu}}$$

Instructive to compare this to result of narrowly distributed case (thin tail), to appreciate the explosive nature of $n^{1/\mu}$:

(•) x exponential : $p(x) = \gamma e^{-\gamma x}$, $x > 0$: $e^{-\gamma M_n^*} \simeq \frac{1}{n}$

$M_n^* \propto \log n$, slow growth with n

(•) x Gaussian

$M_n^* \propto \sqrt{\log n}$, even slower (exercise)(*)

(•) x uniform in $[0,1]$

$M_n^* \simeq 1 - \frac{1}{n}$, converges to the edge

(*) Gaussian exercise : take $\mathcal{G}(0,1)$

$$\Pr(X \geq x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \frac{t}{t} e^{-t^2/2} \, dt$$

by parts $\left[-\frac{1}{t} e^{-t^2/2}\right]_x^{\infty} + \int_x^{\infty} \left(-\frac{1}{t^2}\right) e^{-t^2/2} \, dt = \frac{e^{-x^2/2}}{x} + o\left(\frac{e^{-x^2/2}}{x}\right)$

$$\sim \frac{1}{\sqrt{2\pi}} \frac{1}{x} \exp(-x^2/2)$$

$$\frac{1}{n} \propto \frac{1}{M_n^*} \exp\left(-\frac{{M_n^*}^2}{2}\right) \Rightarrow M_n^* \propto \sqrt{\log n}$$

The result $M_n^* \propto n^{1/\mu}$ allows to understand the generalized CLT stated above

$\mu > 2$: $M_n^* \ll \sqrt{n}$ and the largest element of the set $(x_1, \dots, x_n)$ contributes only marginally to sum S_n → democratic regime, s^{td} CLT

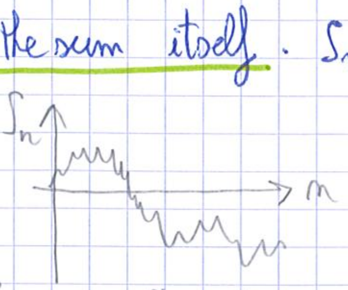
$1 < \mu < 2$: the largest element, $\propto n^{1/\mu}$ thus $\sqrt{n} \ll M_n^* \ll n$

thus it contributes to the fluctuations, but not the ~~mean~~ sum

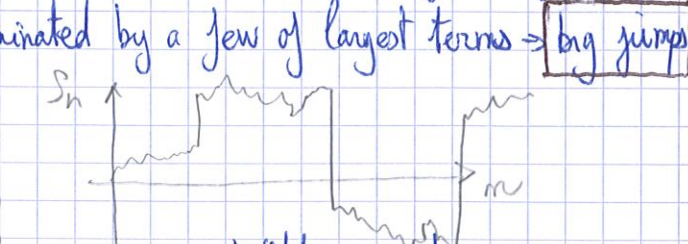
↳ consider here $\frac{S_n - n\langle x \rangle}{n^{1/\mu}}$ for universal behaviour (n independent)

$0 < \mu < 1$: $M_n^* \gg n$ and this largest element contributes a finite fraction of the sum itself. S_n is dominated by a few of largest terms → $\boxed{\text{big jumps}}$

Fig 1 AMIR JSTAT 2020



Summing Gaussian, or $\mu > 2$



Wilder randomness

$\mu < 1$ "Hasard sauvage" (Mandelbrot)

We have a phenomenon of condensation : a finite fraction of the whole sum is explained by a few terms, even when $n \to \infty$ (12)

The 3 types of limiting extreme value statistics (EVS)

We can be more precise, and characterize better the EVS, for all kinds of distrib, broad or not.

$$M_n = a_n + b_n \xi_n \quad ; \quad a_n, b_n \text{ are } p\text{-distribution dependent}$$

When $n \to \infty$, ξ_n can only be of 3 classes, n-independent

1. **Gumbell**, for all $p(x)$ not broad (thin tail) : decay faster than any power law
$$\text{Gumb}(\xi) = \exp\left[e^{-\xi} - \xi\right] \qquad \text{eg, gaussian}$$
2. **Fréchet**, for all $p(x)$ broad with Lévy index μ
$$\text{Fréchet}(\xi) = \xi^{-\mu-1} \exp\left[-\xi^{-\mu}\right]$$
3. **Weibull**, for all $p(x)$ with bounded support (X_{max}) s.t. $p(x) \propto \frac{1}{|X_{max} - x|^{\mu-1}}$ close to the edge. Here $a_n = X_{max}$, $b_n \propto n^{-1/\mu}$
$$\text{Weibull}(\xi) = \xi^{\mu-1} \exp\left(-\xi^{\mu}\right)$$

eg $\mu = 1$ if p is a constant at the edge
$\mu < 1$ for a divergence

4) c) Revisiting the generalized CLT

Knowing $M_n^* \propto n^{1/\mu}$, and with a bit of insight, we can propose an interesting reformulation, that explains also what happens for $\mu = 1$ and $\mu = 2$.

IDEA: after n trials, the region $M_n > M_n^*$ is not sampled / explored.

Truncate $p(x)$ s.t
$$p_{eff}(x) = p(x) \quad ; \quad x \leq M_n^*$$
$$= 0 \quad ; \quad x > M_n^*$$

p_{eff} is almost normalized $\to$ do not correct for this slight effect...

$0 < \mu \leq 1$: $\langle X \rangle = \infty$

$$\langle S_n \rangle = n \langle X \rangle_{eff} \quad ; \quad \langle X_{eff} \rangle = \int_{\varepsilon}^{\infty} x \, p_{eff}(x) \, dx = \int_{\varepsilon}^{M_n^*} x \, p(x) \, dx$$

$$\propto \int_{\varepsilon}^{M_n^*} \frac{x}{x^{1+\mu}} \propto \left(M_n^*\right)^{1-\mu}$$

$$\propto \log M_n^* \quad \text{for } \mu = 1$$

$$\Rightarrow \langle S_n \rangle \propto n \left(n^{1/\mu}\right)^{1-\mu} \propto n^{1/\mu} \quad \mu < 1$$

$$\propto n \log n \quad \text{for } \mu = 1$$

$1 < \mu \leq 2$ $\langle X \rangle$ finite but $\langle X^2 \rangle$ diverges $\Rightarrow$ $V(X)$ diverges

$$V_{eff}(S_n) = V_{eff}\left(\sum_i X_i\right) = V_{eff}\left(\sum_i X_i\right) = n\, V_{eff}(X)$$

$$V_{eff}(X) \propto \int_0^{M_n^*} \frac{x^2}{x^{1+\mu}}\, dx \propto \left(M_n^*\right)^{2-\mu}$$

$$\propto \log M_n^* \quad \text{for } \mu = 2$$

$$\Rightarrow V_{eff}(S_n) \propto n \left(n^{1/\mu}\right)^{2-\mu} \propto n^{2/\mu} \quad \mu < 2$$

$$\propto n \log n \quad \mu = 2$$

6h

4) d) Applications to sub- and super-diffusion

There are many (basic sciences, economy, engineering ...) We address them considering a r-walk:

$$S_n = X_1 + X_2 + \cdots + X_n$$
$$t = \tau_1 + \tau_2 + \cdots + \tau_n$$

the distribution of both the X_i and τ_i can be random $\rightarrow$ behavior of S_t ?

$\longrightarrow$ Take the X_i ~~non-random, or~~ thin tailed, and τ_i distribution broad, index μ

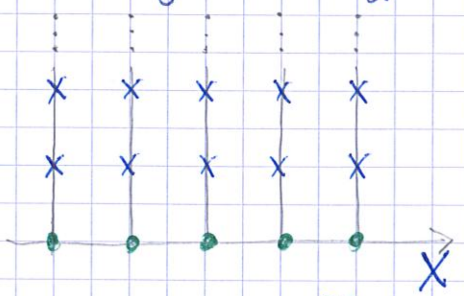
After n steps: $\langle S_n \rangle = 0$ (for simplicity); $\langle S_n^2 \rangle = n\sigma^2$; $\mu < 1$

but $t \propto n^{1/\mu}$

$$\langle S_n^2 \rangle \propto n \propto t^{\mu} \ll t \quad \text{that would hold for regular diffusion}$$

↳ subdiffusion $\boxed{\sqrt{\langle S_n^2 \rangle} \propto t^{\mu/2}} \ll \sqrt{t}$ large t

"Physical realization": diffusion on a comb



The walker jumps towards every neighbors with equal proba (1/2 when on the ⋮ and 1/3 when on the ·–·–·). The return time to nodes • has tail $P(\tau) \propto \tau^{-3/2}$, $\mu = 1/2$.

Thus $X_t \propto t^{1/4}$ SUBDIFFUSION when thus projecting walk onto X axis

$\longrightarrow$ Conversely, if time are not random (thus $t \propto n$), but distribution of jump length is broad with index μ:

$$S_n \propto n^{1/\mu} \quad \text{ie} \quad \boxed{S_t \propto t^{1/\mu} \gg \sqrt{t}}$$

SUPERDIFFUSION for $\mu < 2$

This random walk is called a Lévy Flight.

↳ Lévy flight foraging hypothesis in biology

Physical realization: adsorption/desorption of a molecule diffusing in bulk
↳ on a surface

Mark with • the places where the diffusing A hits the surface (it can get absorbed, then released back in the bulk, this is not essential). The walk projected onto xy is anomalous. Indeed, time τ between 2 hits of the surface is given by return time of 1D r.w (along z, $\perp$ surface) and thus has tail $p(\tau) \propto \tau^{-3/2}$

In time τ, the A explores length $\ell \propto \sqrt{\tau}$ along x and y.

$$p_L(\ell)\, d\ell = p_T(\tau)\, d\tau \propto \frac{1}{\tau^{3/2}}\, d\tau \quad ; \quad \frac{d\tau}{d\ell} \propto \ell$$

$$\Rightarrow p_L(\ell) \propto \ell \frac{1}{\ell^3} \propto 1/\ell^2 \qquad \text{thus } \mu = 1$$

$$\Rightarrow X_t \text{ and } Y_t \propto t \qquad (X_n, Y_n \propto n)$$

↳ SUPERDIFFUSION

4) e) The case of correlated variables

The CLT may survive not too strong correlations. Meaning?

Take $\langle X_i \rangle = 0$ (not restrictive) and consider correlation function

$$C(n) = \langle X_i X_{i+n} \rangle$$, assuming invariance by translation (stationnarity)

Def A problem is stationnary if $p(x_1, x_2, \dots, x_n) = p(x_{1+m}, x_{2+m}, \dots x_{n+m})$ for all m. Think of the index as referring to a position, or to a time

Then $p_1(x_1) = p_1(x_2) = p_1(x_n)$: 1 pt function indep of "time"/"position"

$p_2(x_1, x_{20}) = p_2(x_2, x_{21}) = \dots$ depends only on time/pos difference

→ If $\sum_{j=1}^{\infty} C(j) < +\infty$: finite range correlations

A finite number of neighbors, N_{corr}, is correlated with a given variable

$$N_{corr} = \frac{1}{C(0)} \sum_{j=1}^{\infty} C(j)$$

Group variables into **blocks**, $\frac{n}{N_{corr}}$ in total

$$\underbrace{\times\times\times\times\times\times}_{N_{corr}}\ \underbrace{\times\times\times\times\times\times}_{N_{corr}}\ \times\times$$

$\sum_{block 1} X_i = \hat{X}_1 \qquad \sum X_i = \hat{X}_2$

$$S_n = \sum_{i=1}^{n} X_i = \sum_{j=1}^{n/N_{corr}} \hat{X}_j$$ independent

The scale of each $\hat{X}_i$, sum in block i, is N_{con}

Thus $S_n \propto \sqrt{\frac{n}{N_{con}}}\, N_{con} \propto \sqrt{n N_{con}}$ // $\sqrt{t}\,\ell$

This is the same scaling as the CLT: $\sqrt{n}$.

→ Long-range correlations : means $\sum_j C(j)$ diverges. Take $C(j) \propto \frac{1}{j^a}$

Then $0 < a < 1$

$$\sum_{j\geq 1}^{n} C(j) \propto \sum_{j \geq 1}^{n} 1/j^a \propto n^{1-a}$$

$\Rightarrow N_{con} \propto n^{1-a}$ grows with n but not as fast as n itself. Thus, $\nearrow n$, we have more and more boxes

$$S_n^2 \propto n\, N_{con}$$

$$\boxed{S_n^2 \propto n^{2-a}}$$

Argument is very heuristic.

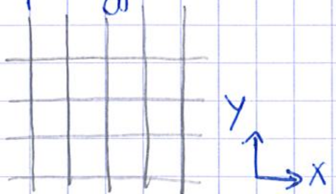
Derivation not rigorous, but result is correct

Correlations are relevant. Note that we recover normal diffusion for $a=1$: for $a > 1$, correlations are short-range ...

Illustration : the MATHERON - DE MARSILY model

Initially proposed to study the diffusion of a tracer in a porous rock : simplest setting where superdiffusion sets in, due to long-range correlations, generated dynamically.

A single p diffuses in a layered medium (say in 2D)



Along Y: standard r.w, $\frac{1}{2}$ up, $\frac{1}{2}$ down

Along X: drift $v(Y)$, the same for all X values, a random function of Y only ($v(Y) = \pm 1$, proba $1/2$, $1/2$), indep of time

The different layers are uncorrelated,

but the motion Y gets correlated due to fact that the same layer may be visited multiple times

This generates a bias in X direction

$$X_{t+1} = X_t + V_{Y_t} \quad ; \quad V_Y = \pm 1, \text{ fixed } \forall t$$

$$Y_{t+1} = \begin{cases} Y_t + 1 & \text{proba } 1/2 \\ Y_t - 1 & \text{" } 1/2 \end{cases}$$

$$X_t = \sum_{i=0}^{t} V_{Y_i}$$

but many of these velocities are correlated (they are even the same!)

In a time t, there are $\sqrt{t}$ sites visited along Y; and among the t terms

Then, typical $X_t \equiv \sqrt{\langle X_t^2 \rangle} \propto \sum_{k=0}^{\sqrt{t}} \sqrt{t}\, V_{Y_k}$ in $\sum_{i=0}^{t} V_{Y_i}$, only $\sqrt{t}$ are $\neq t$

$$\propto \sqrt{t}\,\left(\sqrt{t}\right)^{1/2} \propto t^{3/4}$$ SUPERDIFFUSION

5c) RG view on the CLT, and its generalization   RG = Renormalization Group

Goal is to draw an analogy with theory of phase transitions. We are interested in finding the limiting distribution of a sum of large # of variables, having some $p(x)$

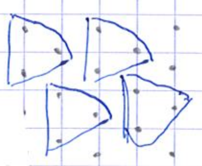
$$S_n = \underbrace{x_1 + x_2}_{\bullet} + \underbrace{x_3 + x_4}_{\bullet} + \dots$$

We proceed recursively: summing pairwise, iterate

We search for a (functionnal) fixed point of this procedure

$\longrightarrow$ reminiscent of RG approach (Wilson, Fisher, … 1970s $\to$ Nobel Wilson 1982) in theory of phase transitions

Idea illustrated on a spin system on triangular lattice for convenience.



Regroup spins in plaquettes of 3 spins. Take majority rule to assign a spin S'_I to every plaquette I

Note that new lattice of plaquette also is triangular (lattice cst $\times \sqrt{3}$)

$$\mathcal{H}(S_1, \dots S_m) = -J \sum_{\langle i,j \rangle} S_i S_j$$

$$\mathcal{H}'(S'_I, S'_J \dots) = ? \quad \text{determined from} \quad \sum_{\{S_i\}} e^{-\beta \mathcal{H}(\{S_i\})} = \sum_{\{S'_I\}} e^{-\beta \mathcal{H}'(\{S'_I\})}$$

Defines a new coupling cst, J'

Iterate $\to$ flow of coupling constants, J, J', J''; all this amounts to zoom out

$T < T_c$: we expect large scale order, hence $J \nearrow$ $J < J' < J'' \to \infty$
$T > T_c$: " disorder, $J > J' > J'' \to 0$
} trivial fixed points

But at $T = T_c$, we expect J to converge to a non trivial fixed point, since system is scale invariant (diverging correlation length).

The analogy below is quite formal, but insightful. From X_1 and X_2, consider $X_1 + X_2$ but we do not know what is the "scale" for analyzing the sum and get a possible fixed point. Thus, take