Central Limit Theorem and Beyond

- CLT ↔ conventional wisdom on random processes, in depth
- Beyond: - Gaussian from CLT, as special case of Levy Stable Distribution
  - Levy Distribution, an Introduction
  - Levy Process - generalizing random walk.

Central Limit Theorem: (De Moivre, Laplace, Gauss)

- Consider a sum of $\sim$-independent random variables, increments:

$$AX_1, AX_2, \ldots, AX_n$$

Let sum $X_n = \sum_{i} AX_i$
- Each \( AX_i \); \( \langle AX_i \rangle = 0 \)
  \[ \langle AX^2_j \rangle = \sigma^2 \]

i.e. Variance of step distribution converges \( \langle AX^2_j \rangle < \infty \).
Only variance required.

- then \( \sum_{j} \sigma^2_j = \sum_{j} \sigma_j^2 \)

CLT \( \Rightarrow \)

\[
\text{Pdf} \ (X_n) \equiv \frac{1}{(2\pi \sigma_n)^{\frac{1}{2}}} \exp(-X_n^2 / 2\sigma_n^2)
\]

\( n \gg 1 \)

\( \therefore \) Pdf sum \( \Rightarrow \) Gaussian

- Key Points / Buried Bodies

\( a1) \ AX_i \) (alike); sum not dominated by few (if so, "intermittency")

\( a2) \) finite variance of step increment \( \langle AX^2_j \rangle < \infty \)
b.1) What of higher moments?

e.g. \[ \langle \Delta x_i^2 \rangle < \infty \neq \langle \Delta x_i^4 \rangle < \infty \]

\[ \Rightarrow \] Large Kurtosis can induce heavy tails (fat tail)

b.2) Quantity ?

Observations:

- CLT states (effectively) that (given conditions satisfied),

\[ X_i \rightarrow \text{statistically distributed} \]

consistent with CLT

then if \( X_i, Y_i \) are series to be summed, which follow CLT

conditions,
then \((a x_i + b) + (a' y_i + b')\)
\[= a'' Z_i + b''\]
is also Gaussian distributed, i.e., follows CLT
("L-clivity")

here: \(a, b, j, a', b'\) all \(\geq 0\) and
not stochastic

In simple terms:

\[\iff\]
Adding two Gaussian distributed series yields a sum which is
Gaussian distributed.

\[\implies\]
CLT \(\Rightarrow\) "Gaussianity modulo conditions, 
is an attractor in function space"

More generally: A \(\square\) of
distributions exist which are
L-stable (L for Paul Levy)

c.o have property that if two series distributed, sum is also distributed similarly.

The Message:

The skewed Gaussian of CLT is merely one particular case of an L-stable distribution, and the only one with finite variance.

Many elements in class of L-stable attractors on function space

Family of allowed distributions is larger than you thought...

To understand: First re-visit CLT
Proving the C. L. T.

General ideas:
- Markov process \( \rightarrow \) Chapman - Kolmogorov
- \( E_{2n} \)
- Convolutions
- Convolution \( \rightarrow \) Product of F.T.
- "Generating" on "characteristic" Function

Point: Fourier Transform of step probability is more significant than probability. (and useful)

So \( C \rightarrow K \) \( E_{2n} \):
- Takes \( x-y \rightarrow x \)

\[
P_N (x) = \sum y \ P_N (y) \ P_N (x | x+y)
\]

Don't expand ----
then, if F.T. and noting F.T. (convolution)

\[ F \{ F.T. \} \quad \text{i.e. Fourier transform of convolution} = \text{product of functions convolved.} \]

then, \( n \) step C-LT:

\[ P_N(x) = P_n(x) \prod_{j \neq n} P_j(x) = \sum_{n=1}^{N} P_n(x) \]

\[ P_n(x) = \sum_{k=1}^{N} e^{ikx} \prod_{j \neq n} P_j(x) \quad \text{applies for identical steps.} \]
Can also define moments:

\[ m_n = \int dx \ x^n \ P(x) \]

\[ \langle x^1 \rangle = m_1 \]
\[ \langle x^2 \rangle = m_2 \]
\[ \langle x^n \rangle = m_n \]

\[ \hat{P}(k) = \sum_{n=0}^{\infty} \left( \frac{-i}{k} \right)^n \ \frac{m_n}{n!} \]

\[ = \int e^{-i k x} \ dx \ P(x) \]
\[ = \int dx \ (1 - i k x + \frac{(k x)^2}{2} + \ldots) \ P(x) \]
\[ = \int dx \ \sum_{n=0}^{\infty} \left( \frac{-i}{k} \right)^n \ \frac{m_n}{n!} \ x^n \ P(x) \]

\[ \Rightarrow m_n = \int \frac{\partial^n}{\partial (ik)^n} P(x) \]

Useful identity:  \( n^{th} \) moment is \( n^{th} \) derivative of generating function.
\[ \hat{\rho} (k) = 1 - i M k - \frac{1}{2} k^2 M + ... \\
\text{easily generalized to higher dimensions.} \]

3. Cumulants

- i.e. nonlinear combinations of moments

\[ \psi (k) \equiv \ln \hat{\rho} (k) \]

\[ \rho (x) = \int \frac{e^{ikx}}{(2\pi)} \hat{\rho} (k) \frac{dk}{2\pi} \]

\[ = \int e^{i[kx + \psi (k)]} \frac{dk}{2\pi} \]

Expand:

\[ \psi (k) = -i C_1 k + \frac{1}{2} c_2 k^2 + ... \]

(Closed in terms of cumulants)

\[ C_1 = m_1 \]

\[ c_2 = m_2 - m_1^2 = \sigma^2 \]

et. al.
If exist, one has $F_{\omega m}$ moments:

$\int_{C_n} (\omega - \omega_m)^n = \frac{\delta^n}{\omega_{\xi_1}, \omega_{\xi_2}, \ldots, \omega_{\xi_n}}$

Now assuming independent, independently distributed (i.i.d.) steps, cumulants additive:

$\hat{r}_N(x) = \int e^{i k x} \, d k \, \hat{r}_N(k)$

$= \int e^{i k x} \frac{d k}{2 \pi} (e^{i k x})^N$

$= \int \frac{d k}{2 \pi} e^{i k x} (e^{i k x})^N$

$= \int \frac{d k}{2 \pi} e^{i k x} e^{Nkx}$

$= \hat{r}_N(x) = N \hat{r}_N(x)$
So, for C.L.T.: Consider \( N \to \infty \)

(Casny mptotic \( \infty \))

\[
\begin{align*}
\hat{p}_N(x) &= \sum_{n=1}^{\infty} \int \frac{\hat{p}_n(k)}{n!} e^{ikx} dk \\
&= \int e^{ikx} dk \left( \frac{\hat{p}(k)}{2\pi} \right)^n \\
&= \int e^{ikx} \frac{dk}{2\pi} e^{-N \psi(k)}
\end{align*}
\]

i.e. additivity: \( \hat{p}_N = N \psi(k) \)

\[\psi(k) = -i c \kappa - k^2 \xi \]

\[\hat{p}_N(x) = \int \sqrt{2\pi} \frac{e^{ikx}}{\sqrt{\kappa}} e^{-N \psi(k)}
\]

For \( N \to \infty \) only the region near \( \kappa = 0 \) contributed (Laplace's Method)

Only low order cumulants contribute to determine \( \hat{p}_N(x) \)

C.N.B. Fundamental reasoning for truncating (Moments - Moyal)
\[ P(x,t) = \frac{1}{(\sqrt{2\pi t})^2} \exp\left[-\frac{x^2}{2t}\right] \]
etc.

\[ \Rightarrow \text{C.L.T.} \]

A few points:

- No questions asked about higher moments for \( N \to \infty \).

- These need not be well behaved and can induce Fat tails.

\[ \therefore P(x) = \frac{1}{1 + x^2} \]

\[ \text{hence } \langle x^2 \rangle \to \infty, \text{ so C.L.T. not apply} \]

but
\[ P(x) = \frac{1}{\sqrt{2\pi}} (1 + x^4) \]

\[ \langle x^2 \rangle < \infty \]

meets C.L.T. criterion, but kurtosis diverges

\[ \Rightarrow \text{Fat Tail} \]
- Can show,

- Gaussian eroded (fat tail, + probability conserved = erode central)
  = erode central Gaussian)

- large $x$
  (how large is "large"?)

$P_N(x) \sim N/A/4^x$

(power law, not Gaussian)

N.B. - Refs: Chandrasekher Review
  Hikos et al.
  Hughes, B.D.
  or any book...

- M.I.T. OCW 18.366
  ("Random Walks and Diffusion")

- Physics 235, Spring '19
  (Note: write-ups, Supplementary)

N.B. Issue of Fat Tail behavior within
  CLT is good paper topic.
- Levy Distributions

- observe: A property of diffusion \( \rightarrow \) Self-Similarity

\[
D = \frac{\langle dx^2 \rangle}{dt}
\]

\[
\begin{align*}
\Delta x & = x \Delta x', \\
\Delta t & = \Delta t'
\end{align*}
\]

\[
D' = \alpha^{-2} \langle \Delta x'^2 \rangle = 1
\]

\[
\frac{\partial^2}{\partial t^2}
\]

\[
\begin{cases}
\alpha = \chi^2
\end{cases}
\]

- What is the class of self-similar distributions which are L-stable and normalizable?

Now, \( x_i \rightarrow \) random variable

\[
X_N = \sum_{i=1}^{N} X_i
\]

\[
\text{generating characteristic function:}
\]

\[
\hat{P}(k) = [P(k)]^N
\]
Rescale: \[
\begin{align*}
Z_N &= X_N / a_N \\
\text{Pdf} (Z_N) &= F_0 (X) / a_N \\
X &= X_N / a_N
\end{align*}
\]

\[\hat{F}_N (a_N k) = \hat{F}_0 (k)\]

Now seek attractors in function space, so:

\[F_N (k) \rightarrow \hat{F}(k)\]

\[\Lambda \rightarrow \infty\]

Let \( \lim_{m \to \infty} \frac{a_{nm}}{am} = c_n. \) (coarsest Loziability)

Then have condition for function \( \hat{F}(k) \) as limiting case:

\[\hat{F}(k, c_n) = \left[ \hat{F}(k) \right]^n\]

\[\text{scale} \rightarrow \text{self-similarity}\]
So need solve

\[ \psi (k \mu(x)) = (F(k))^{x} \]

\[ \psi (k \mu(x)) = \lambda \psi (k) \]

with \( \mu (x=1) = 1 \).

\[ \kappa \frac{d}{dk} \psi (k \mu(x)) = \psi (k) \]

\[ \kappa \mu(x) \psi'(x) = \psi' \]

\[ \frac{d \psi}{dk} = \frac{\psi}{\mu(x) \kappa} \]

\[ \text{Power law for } \psi \]

\[ \psi (k) = \begin{cases} v_1 \mu_1^\kappa, & k \geq 0 \\ v_2 \mu_1^\kappa, & k < 0 \end{cases} \]
Can show or more detail: (Hutches)

\[ \hat{F}(\ell) = \exp \left[ -\frac{1}{2} \ell \beta^2 \right] \]

\[ \text{levy distribution} \]

\[ L_\alpha(a, k) = \hat{F}(k) = \exp(-a |k|^\alpha) \]

\[ \alpha = 2 \Rightarrow \hat{F}(k) = \exp(-ak^2) \rightarrow \mathcal{L}_2 \]

\[ \alpha = 2 \text{ is self-similar attractor} \]

\[ L\text{-stable with normalizable} \]

\[ \text{2nd moment (CC.LT. case) (only)} \]

Can show \( \alpha = 2 \) is max. \( \alpha \).

\[ \alpha = 1 \Rightarrow \hat{F}(k) = C^{-\alpha} |k|^{-\alpha} \rightarrow \text{Cauchy Lorenz-Zygm} \]

\[ P(x) = \frac{1}{\sqrt{\alpha^2 + x^2}} \]