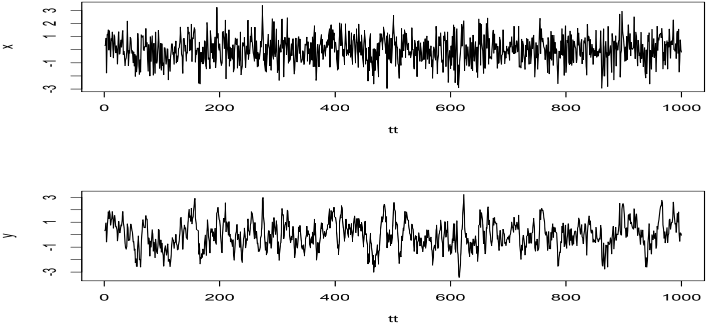
Beyond the color of the noise: what is "memory" in random phenomena ?

Gennady Samorodnitsky
Cornell University

September 19, 2014

Randomness means lack of pattern or predictability in events

according to **Wikipedia**



However: certain different patterns are present on the two plots.

- The two plots are of two stationary stochastic processes with the same marginals.
- The second one has memory, while the first one does not.

Traditionally, in probability the notion of memory applies to stationary stochastic processes $(X_n,\ n = 0, 1, 2, \ldots)$: for every $h = 1, 2, \ldots$

$$(X_n,\ n = 0, 1, 2, \ldots) \stackrel{\mathrm{d}}{=} (X_{n+h},\ n = 0, 1, 2, \ldots).$$

"The memory" in a stochastic process: how observations far away in time affect each other.

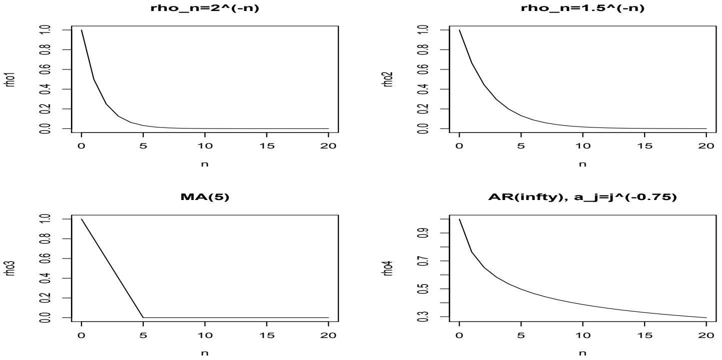
How does one measure memory?

It is obvious: use correlations!

Let

$$\rho_n = \mathrm{Corr}\big(X_k, X_{k+n}\big), \ \ n = 0, 1, 2, \ldots .$$

Four different correlation functions



What do we see in these plots?

Covariances and correlations of a second-order stationary process can be expressed through the spectral measure of the process:

$$\rho_n = \frac{1}{\mathrm{Var}\, X_0} \int_{(-\pi,\pi]} e^{inx}\, F(dx), \;\; n = 0, 1, 2, \dots ;$$

F is a finite symmetric measure on $(-\pi, \pi]$. If F has a density with respect to the Lebesgue measure on $(-\pi, \pi]$,

the density f is called the power spectral density of the process.

One can view the process as the sum of waves of different frequencies with random and uncorrelated weights:

$$X_n = \int_{(-\pi,\pi]} e^{inx}\, M(dx), \;\; n = 0, 1, 2, \dots\,,$$

M a random measure governed by the spectral measure (density).

Such a process is also called **a noise**.

- If the spectral density is constant, the noise is white.
- If some frequencies have a larger weight than some other frequencies, the noise is colored.
- The common colors of the noise: pink, brown, blue, violet, grey.
- The different colors describe which frequencies are preferred and by how much.

Long range dependence (long memory)

Long memory in a stochastic process:

when "the common wisdom goes wrong", for example:

- The "square root of the sample size" rule is no longer valid;
- Objects that used to be approximately normal are no longer so.

In the second-order language:

- either slowly decaying correlations

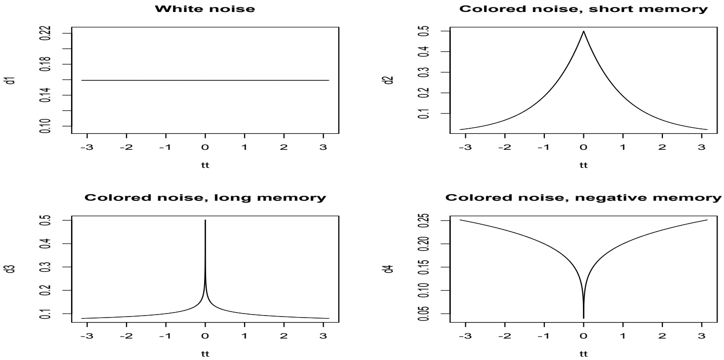
$$\rho_n \sim n^{-d} L_n, \ n \to \infty \,,$$

$0 < d < 1$, L_n is slowly varying,

- or spectral density has a pole at zero

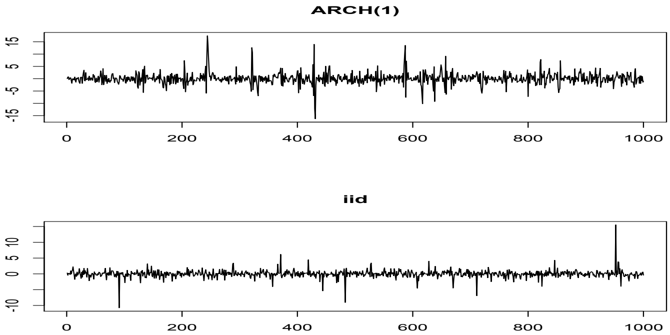
$$f(x) \sim x^{d-1} L(1/x), \ x \downarrow 0 \,.$$

Four different spectral densities



How much information is in the color of the noise?

Very different processes, the same color



Both processes are white noises, with equally heavy tails

Conclusions:

- the second order characteristics do not provide enough information about the length of the memory;
- long range dependence is far from being determined by correlations;
- with sufficiently heavy tails correlations do not even exist;
- long memory as a phenomenon is a phase transition.

Examples of long memory as a phase transition

1 Suppose that $\big(X_n,\ n = 0, 1, 2, \ldots\big)$ is a stationary stochastic process with a finite variance.

Let

$$S_n = \sum_{i=1}^{n} X_i,\ \ n = 1, 2, \ldots .$$

We may say that the process has short memory if, as $n \to \infty$,

$$\frac{S_n - nEX_0}{n^{1/2}} \Rightarrow N(0, \sigma^2),\quad 0 < \sigma^2 < \infty\,.$$

We may require that

$$\left(\frac{S_{[nt]} - [nt]EX_0}{n^{1/2}},\ t \geq 0\right) \Rightarrow \left(\sigma B(t),\ t \geq 0\right)$$

in $D[0, \infty)$ in the Skorohod topology.

The limit $\left(B(t),\ t \geq 0\right)$ is the Brownian motion. It is

- a Gaussian process,
- has stationary increments,
- is self-similar with $H = 1/2$.

We may say that the process has long memory if:

- as $n \to \infty$, the normalization $n^{1/2}$ is wrong:

$$\frac{S_n - nEX_0}{a_n} \text{ has a finite nonzero limit;}$$

- when $a_n >> n^{1/2}$, the memory is long and positive;
- when $a_n << n^{1/2}$, the memory is long and negative (also: medium memory).

Under long memory the limit $\mathbf{Y}$ in

$$\left(\frac{S_{[nt]} - [nt]EX_0}{a_n},\ t \geq 0 \right) \Rightarrow \left(Y(t),\ t \geq 0 \right)$$

is a self-similar process with stationary increments, but not a Brownian motion. It could be:

- a Fractional Brownian motion (a Gaussian process);
- a process in a higher order Wiener chaos:

$$Y(t) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} Q_t(x_1, \ldots, x_k)\, dW(x_1) \ldots dW(x_k),\ \ t \geq 0\,.$$

2 Suppose that $\mathbf{X} = \big(X_n,\ n = 0, 1, 2, \dots\big)$ is a stationary stochastic process with infinite variance.

For simplicity: asssume the process to be symmetric: $-\mathbf{X} \stackrel{\mathrm{d}}{=} \mathbf{X}$.

Assume that the observations have regularly varying tails:

$$P(|X_n| > x) = x^{-\alpha} L(x),\ x > 0,\ 0 < \alpha < 2\,,$$

where L is a slowly varying at infinity function.

If the memory is short, then

$$\left(\frac{S_{[nt]}}{a_n},\ t \geq 0 \right) \Rightarrow \left(L_\alpha(t),\ t \geq 0 \right);$$

- $a_n = n^{1/\alpha} L_1(n)$, L_1 is slowly varying;
- L_α is a symmetric α-stable Lévy motion;
- L_α is self-similar with $H = 1/\alpha$.

If the memory is long, one starts getting, as limits, processes such as Linear Fractional Stable Motions:

$$Y(t) = \int_{\mathbb{R}} \left(|t-x|^{H-1/\alpha} - |x|^{H-1/\alpha} \right) L_\alpha(dx) \ \ t \geq 0 \, .$$

- The range of H: $0 < H < 1$;
- the process $\mathbf{Y}$ is H-self-similar, and has stationary increments.
- The memory can be even longer.

The stationary process X

We consider infinitely divisible processes of the form

$$X_n = \int_E f_n(x) dM(x), \quad n = 1, 2, \ldots .$$

- M is a homogeneous symmetric infinitely divisible random measure on a $(E, \mathcal{E})$.
- M has an infinite, σ-finite, control measure μ and local Lévy measure ρ: for every $A \in \mathcal{E}$ with $\mu(A) < \infty$, $u \in \mathbb{R}$,

$$Ee^{iuM(A)} = \exp\left\{ -\mu(A) \int_{\mathbb{R}} \left(1 - \cos(ux)\right) \rho(dx) \right\} .$$

The functions f_n, $n = 1, 2, \ldots$ are deterministic functions of the form

$$f_n(x) = f \circ T^n(x) = f\big(T^n x\big),\ x \in E,\ n = 1, 2, \ldots :$$

- $f : E \to \mathbb{R}$ is a measurable function, satisfying certain integrability assumptions;
- $T : E \to E$ a measurable map preserving measure μ.

We assume that the local Lévy measure ρ has a regularly varying tail with index $-\alpha$, $0 < \alpha < 2$:

$$\rho(\cdot, \infty) \in RV_{-\alpha} \text{ at infinity.}$$

With a proper integrability assumption on the function f:

the process X has regularly varying finite-dimensional distributions, with the same tail exponent $-\alpha$.

The key assumption:

- the map T is conservative and pointwise dual ergodic: there is a sequence of positive constants $a_n \to \infty$ such that

$$\frac{1}{a_n}\sum_{k=1}^{n} \widehat{T}^k f \to \int_E f \, d\mu \quad \text{a.e.}$$

  for every $f \in L^1(\mu)$.
- The dual operator $\widehat{T}$ satisfies the relation

$$\int_E \widehat{T} f \cdot g \, d\mu = \int_E f \cdot g \circ T \, d\mu$$

  for $f \in L^1(\mu)$, $g \in L^\infty(\mu)$.

Theorem

Assume that the normalizing sequence (a_n) in the pointwise dual ergodicity is regularly varying with exponent $0 < \beta < 1$ and that $\mu(f) = \int f \, d\mu \neq 0$. Then for some sequence (c_n) that is regularly varying with exponent $\beta + (1-\beta)/\alpha$,

$$\frac{1}{c_n} \sum_{k=1}^{\lfloor n \cdot \rfloor} X_k \Rightarrow |\mu(f)| Y_{\alpha,\beta} \quad \text{in } D[0,\infty) \, .$$

The limiting process

Let $0 < \beta < 1$. We start with inverse process

$$M_\beta(t) = S_\beta^{\leftarrow}(t) = \inf\{u \geq 0 : \ S_\beta(u) \geq t\}, \ t \geq 0 \,.$$

- $\big(S_\beta(t), t \geq 0\big)$ is a (strictly) β-stable subordinator.
- $\big(M_\beta(t), \ t \geq 0\big)$ is called *the Mittag-Leffler process*.

The Mittag-Leffler process has a continuous and non-decreasing version.

- It is self-similar with exponent β.
- Its increments are neither stationary nor independent.
- All of its moments are finite.

$$E \exp\{\theta M_\beta(t)\} = \sum_{n=0}^{\infty} \frac{(\theta t^\beta)^n}{\Gamma(1+n\beta)}, \quad \theta \in \mathbb{R}.$$

Define

$$Y_{\alpha,\beta}(t) = \int_{\Omega' \times [0,\infty)} M_\beta\big((t-x)_+, \omega'\big) dZ_{\alpha,\beta}(\omega', x), \quad t \geq 0.$$

- $Z_{\alpha,\beta}$ is a SαS random measure on $\Omega' \times [0, \infty)$ with control measure $P' \times \nu$.
- ν a measure on $[0, \infty)$ given by $\nu(dx) = (1-\beta)x^{-\beta}\, dx$.
- M_β is a Mittag-Leffler process defined on $(\Omega', \mathcal{F}', P')$.

Conclusions

- The length of memory should not be measured by correlations or a similar simple measure.
- Deeper features of the process affect the limiting distributions of the partial sums, partial maxima, etc.
- In each given application it is important to use a model that "fits", and not only the spectrum of the noise.