Intro to Intermittency (Lectures 9-11)

Lectures by Pat Diamond
Notes by Kenneth Gage

What Is Intermittency?

Intermittency in either space or time is described by a probability distribution with widely spaced patches of high probability concentration. In other words, intermittent events have local probabilities much higher than the mean $P(x_i) \gg \bar{P}$.

Intermittency is what Mandelbrot refers to as "wild" randomness as opposed to the "mild" randomness of Gaussian statistics[1] — the realm of the Central Limit Theorem and Law of Large Numbers.

Intermittency is related to multiplicative noise processes, growth in high order moments, and fractal geometries, each of which is explained in these notes.

Additive vs Multiplicative Processes

In order to compare to what we know already, let's take a look back at the Central Limit Theorem: Suppose you have a sequence of independent, identically distributed random variables, each with a mean μ and variance σ^2 ($\bar{x}_i = \mu$, $\langle (x_i - \bar{x}_i)^2 \rangle = \sigma^2$). Then the sum

$$\frac{1}{n}\sum_{i=1}^{n} x_i - \mu$$

is normally distributed according to a Gaussian of width (standard deviation) $\sqrt{n}\sigma$. This is proved by showing using the Fourier transform of the distribution function and the I.I.D. property of x_i[2].

This theorem is the basis for the Langevin equation[3] and additive noise

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\mu v + \frac{\tilde{f}}{m},$$

the Fluctuation-Dissipation Theorem

$$T \cong \frac{\langle \tilde{f}^2 \rangle \tau_{ac}}{\mu},$$

and the Fokker-Planck Equation. The important takeaways from this method are

- the process is additive,
- steps are uncorrelated,
- steps are identically distributed (none are special),
- and the variance exists (no "fat tails").

On the other hand, we have multiplicative processes. Consider this time a product $X \equiv \prod_{i=1}^{N} x_i$ where each x_i can take a value of either 0 or 2 with equal probability. Now the value of X must be zero unless all $x_i = 2$, in which case, $X = 2^N$. The probability of getting this large result is $P(X = 2^N) = 2^{-N}$. Note that N can be thought of as a stand-in for time or a number of steps.

While the spikiness of this toy model is contrived, it is still useful to look at moments of the distribution for some insights. The ensemble average over possible outcomes is fairly simple:

$$\begin{aligned}\langle X \rangle &= \frac{\sum_j X_j}{2^N} \\ &= \frac{0 + 0 + ... + 0 + 2^N}{2^N} \\ &= 1.\end{aligned}$$

The second moment is

$$\begin{aligned}\langle X^2 \rangle &= \frac{\sum_j X_j^2}{2^N} \\ &= \frac{0^2 + 0^2 + ... + 0^2 + (2^N)^2}{2^N} \\ &= 2^N,\end{aligned}$$

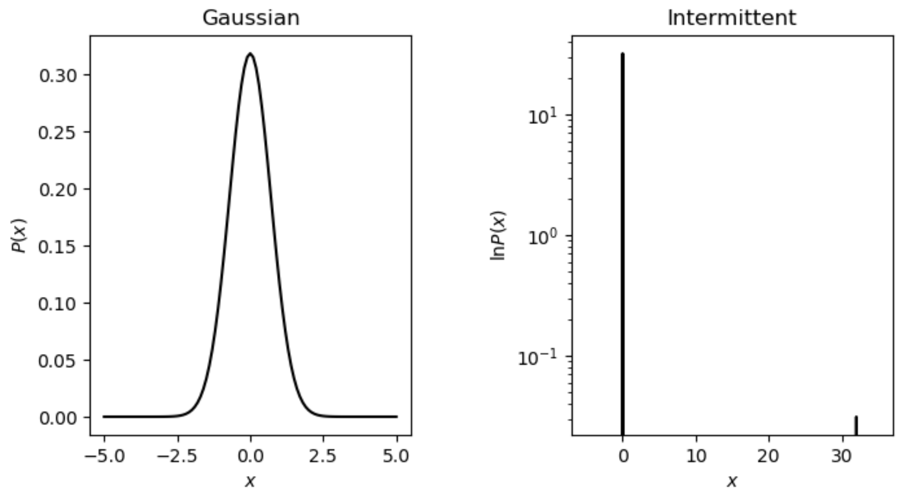

Figure 1: The PDF of a Gaussian distribution ($\mu = 0$, $\sigma = 1/\sqrt{2}$) and the log PDF of our toy intermittency model ($N = 5$).

and the pattern for higher moments (made clear by the lack of nonzero terms) must be $\langle X^p \rangle = 2^{(p-1)N}$, showing higher moments grow exponentially:

$$\gamma_p = \frac{\log_2 \langle X^p \rangle}{N} = p - 1.$$

The structure of this distribution (Fig) is very different from the smooth Gaussian. It is spiky, and the spike at 2^N can be seen as a fat tail, where the Gaussian's is thin. These tails are important for driving the larger moments of intermittent processes to grow (for Gaussians, they ensure the convergence of higher moments).

General takeaways should be:

- intermittent quantities come from products of random numbers (not sums),
- higher moments are important (growth or a lack of convergence),
- and PDFs are concentrated, with heavy tails.

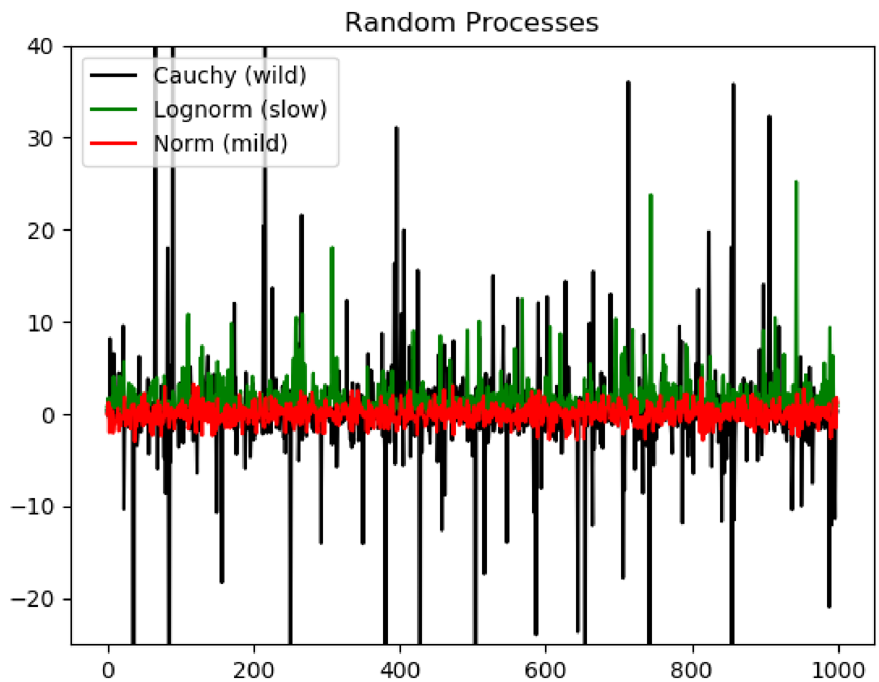

Figure 2: 1000 random samples taken from three distributions. Each of these has a different "level" of randomness or spikiness based on the tail of the underlying distribution generating the data.

Multiplicative Processes and the Log-normal Distribution

These multiplicative processes, or "slow" randomness[1], are related to the log-normal distribution. As a more general version of the previous example, suppose you have a series of positive x_i randomly distributed variable and $X = \prod_i x_i$. The Central Limit Theorem does not apply to X, but it does apply to $\ln X = \sum_i \ln x_i$, which is additive. If $\ln x_i$ is distributed about zero — x_i approximately distributed about one — then the same process as above will find that the probability of getting a value of $\ln X$ is normally distributed:

$$P(\ln X) \sim \exp\left[-\frac{(\ln X)^2}{2N\sigma^2}\right],$$

with the more general case being

$$P(\ln X) \sim \exp\left[-\frac{(\ln X - \mu)^2}{2N\sigma^2}\right].$$

This makes X log-normally distributed

$$P(X) \sim \frac{1}{X}\exp\left[-\frac{(\ln X - \mu)^2}{2N\sigma^2}\right].$$

Since the width of $P(\ln X)$ scales with $N^{1/2}$, we can approximate $\ln X$ by $N^{1/2}\eta$ with η being a normally distributed number with μ and σ^2 as mean and variance respectively. The relationship for the variable itself, however, is $X \sim \exp\big(N^{1/2}\eta\big)$. For large values of N, the value of X is expected to be either very large or very small depending on the sign of η.

To see the behavior of moments of X, we begin with the mean:

$$\begin{aligned}
\langle X \rangle &= \int X(\eta) P(\eta)\, \mathrm{d}\eta \\
&= \int \exp\big(N^{1/2}\eta\big) \exp\left(-\frac{(\eta-\mu)^2}{2\sigma^2}\right) \mathrm{d}\eta \\
&\sim \exp\big(N\sigma^2/2\big),
\end{aligned}$$

where the last step comes from completing the square in the previous exponential. The same exercise can be carried out for higher order moments, giving

$$\langle X^p \rangle \sim \exp\left(\frac{p^2 N\sigma^2}{2}\right).$$

Exponential growth in the moments is then

$$\begin{aligned}
\gamma_p &= \lim_{N\to\infty} \frac{\ln\langle X^p \rangle}{N} \\
&\sim \frac{p^2\sigma^2}{2}.
\end{aligned}$$

Evolution of Randomness

In all of the examples that we've looked at so far, we have only talked about number of samples as being a scale for the system, but we can also view it as an analog for how randomness is affected by time. Let's look from the perspective of random noise: In an additive process

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = \epsilon(t),$$

we can see how the noise leads to diffusion if ϵ is delta correlated

$$\langle \epsilon(t_1)\epsilon(t_2)\rangle = \sigma^2 \tau_{ac} \delta(t_1 - t_2).$$

Here the variance of ψ gives

$$\begin{aligned} \langle \psi^2 \rangle &\sim \int \mathrm{d}t_1 \epsilon(t_1) \int \mathrm{d}t_2\, \epsilon(t_2) \\ &\sim \sigma^2 \tau_{ac} t. \end{aligned}$$

Alternatively, the Central Limit Theorem can be applied since we assume that each time step τ_{ac} is uncorrelated[4].

$$\begin{aligned} \psi &= \int \epsilon(t)\, \mathrm{d}t \\ &= \int_0^{\tau_{ac}} \epsilon(t)\, \mathrm{d}t + \int_{\tau_{ac}}^{2\tau_{ac}} \epsilon(t)\, \mathrm{d}t + ... \\ &= \langle \epsilon \rangle t + \tau_{ac}\sigma \left[\frac{t}{\tau_{ac}}\right]^{1/2} \eta, \end{aligned}$$

with η randomly distributed on a standard Gaussian. If we take the noise to have an average of zero, we get

$$\begin{aligned} \langle \psi^2 \rangle &= \sigma^2 \tau_{ac} t \langle \eta^2 \rangle \\ &\sim \sigma^2 \tau_{ac} t. \end{aligned}$$

The multiplicative process is somewhat different:

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = \epsilon(t)\psi.$$

For this case, Quasilinear Theory can actually get the lowest order moment correct. Assuming $\psi = \langle\psi\rangle + \tilde{\psi}$, we can split the ODE in two.

$$\frac{\mathrm{d}\langle\psi\rangle}{\mathrm{d}t} = \langle\epsilon\tilde{\psi}\rangle$$

$$\frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}t} = \epsilon\langle\psi\rangle$$

The fluctuation equation shows that

$$\tilde{\psi} = \int \epsilon\langle\psi\rangle \,\mathrm{d}t\,,$$

so the ODE for the mean can be written as

$$\begin{aligned}\frac{\mathrm{d}\langle\psi\rangle}{\mathrm{d}t} &= \langle\epsilon\int \mathrm{d}t\,\epsilon\langle\psi\rangle\rangle\\ &\sim \sigma^2\tau_{ac}\langle\psi\rangle.\end{aligned}$$

It follows that the mean is

$$\langle\psi\rangle \sim \exp\bigl(\sigma^2\tau_{ac}t\bigr).$$

While this works for the lowest order moment, everything else requires a different approach. Luckily enough, that is to simply work with the logarithm.

$$\frac{\mathrm{d}\ln\psi}{\mathrm{d}t} = \epsilon(t)$$

The logarithm behaves exactly as the additive process would, but now we have

$$\psi(t) \sim \psi_0 \exp\left(\int \epsilon\,\mathrm{d}t\right).$$

Using the previous result

$$\ln\psi = \langle\epsilon\rangle t + \sigma\sqrt{\tau_{ac}t}\eta,$$

we find that for noise with zero mean,

$$\psi \sim \exp\bigl(\sigma\sqrt{\tau_{ac}t}\eta\bigr).$$

As stated earlier, the QLT average holds

$$\begin{aligned}\langle\psi\rangle &= \int \psi(\eta)P(\eta)\,\mathrm{d}\eta \\ &\sim \int \exp\bigl(\sigma\sqrt{\tau_{ac}t}\eta\bigr)\exp\bigl(\eta^2/2\bigr)\,\mathrm{d}\eta \\ &\sim \exp\bigl(\sigma^2\tau_{ac}t\bigr).\end{aligned}$$

As before with the log-normal distribution, $\gamma_p \sim p^2\sigma^2\tau_{ac}$.

To summarize:

- ψ grows exponentially with time ($\sim \exp\bigl(t^{1/2}\bigr)$),
- fluctuations also grow exponentially,
- QLT can get the mean right, but not structure of higher moments,
- large higher order moments are a signature of "slow" intermittency,
- and intermittency comes from rare, but *intense*, peaks in random behavior.

Intermittent Examples in Random Media

IF we want to look at particle statistics — say the density n of $u(\vec{x},\omega)$ at temperature T — we can see how randomness in one variable affects another. With u distributed according to a Gaussian and ω labeling an ensemble, we can look at the probability of a density

$$P(n(u)) = n(u)P(u).$$

Using the Boltzmann distribution $n = n_0\exp(-u/T)$, we get

$$P(n) = \frac{n_0}{\sqrt{2\pi}\sigma}\exp\Bigl(-\frac{u}{T}\Bigr)\exp\biggl(-\frac{u^2}{2\sigma^2}\biggr).$$

The usual trick with completing the square shows that not only is $P_{max} \sim n_0/\sigma\exp(\sigma^2/2T)$, but the moments go as

$$\langle n^p\rangle^{1/p} = n_0\exp\biggl(\frac{p\sigma^2}{2T}\biggr).$$

Once again we have the familiar signs of exponential growth in the higher order moments, and we see that even though u is normally distributed, the statistics of n have large deviations from the Gaussian.

For a more meaningful example, let's look at the something a type of Reaction-Diffusion Equation:

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = D\nabla^2\psi + u(\vec{x},\omega)\psi.$$

Here we are using u as a stochastic potential with limited range ($\langle u(x)u(x')\rangle \to 0$ for $|x - x'| > l$)

Note that this equation is similar to Fisher's Equation ($[\partial_t - D\nabla^2 - ru(1-u)]\psi = 0$) with only the linear term, so it has some relevance. It is also similar to the Schrodinger Equation, but without imaginary time. As in quantum mechanics, we will use a (Wiener) path integral to tackle this differential equation. Redefining $\mathrm{d}\psi/\mathrm{d}t \equiv H\psi = (H_0 + u)\psi$, we formulate our solution as

$$\psi = \psi_0 \exp\Big(\int H_0 \,\mathrm{d}t\Big) \exp\Big(\int u \,\mathrm{d}t\Big).$$

If we average over diffusion trajectories, we find

$$\langle\psi\rangle = M_X[\psi_0 \exp\Big(\int u \,\mathrm{d}t\Big)],$$

with M_x defined as the trajectory weighted average. This is taken by using the independent nature of steps

$$P(t_1, t_2 - t_1, ...; x_1, x_2 - x_1, ...) = \prod_{j=1}^{n} \frac{1}{[2\pi(t_j - t_{j-1})]^{1/2}} \exp\Big[-\frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})}\Big].$$

It should be remembered that, being Brownian motion, Wiener paths have several properties. The paths $w_t(\omega)$

- are equivalent to locations $x(t,\omega)$,
- start at the origin ($w_0 = 0$),
- are Gaussian distributed,
- have zero mean and finite variance between time steps ($w_{t+\tau} - w_t$),

- and they are continuous but not differentiable.

The lack of differentiability comes from the fact that $\Delta w \sim \sqrt{\Delta t}$, so the limit of $\Delta t \to 0$ of $\Delta w/\Delta t$ isn't finite.

While the averaging can be complicated, the dominant contribution comes from the area around the maximum potential. If the peak has radius R and your grid sizing is l, you should get

$$(\frac{R}{l})^3 P \sim 1,$$

with a probability of approximately $\exp(u_0^2/2\sigma^2)$ for the region. A bit of algebra finds that $u_0 \sim [6\sigma^2 \ln(R/l)]^{1/2}$. If we remember that the region of interest grows with time for the diffusive trajectories ($R \sim \sqrt{Dt}$), we get a solution of

$$\psi \sim \exp\{t \left[3\sigma^2 \ln\left(\frac{Dt}{l^2}\right)\right]^{1/2}\}.$$

From this, we can draw several conclusions:

- the solution is weakly super-exponential,
- the dominant contributions of u in the Wiener path averaging determine the evolution,
- and path integrals can be useful for intermittency.

Spacial Intermittency

The way turbulence was approached by Kolmogorov[5] was to assume a uniform distribution of dissipation; however, real dissipation is distributed in patches with varying intensity, and the whole of space isn't filled. In order to characterize the geometry and dimensionality of the structures created, we must look toward fractal models like the β model.

To start thinking about fractional dimensionality, we can use the box-counting dimension. Consider an object in usual Cartesian space: if it takes $N(\epsilon)$ cubes of size ϵ to completely cover the object, we can define its box-counting dimension as

$$D_0 = \lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln 1/\epsilon}.$$

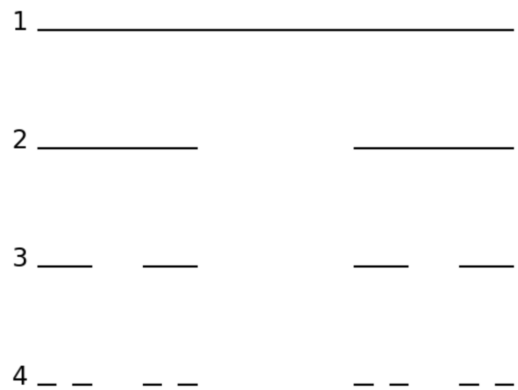

Figure 3: Several consecutive steps in the Middle Third Cantor Set.

As an example, to cover a finite number of points p, you need p boxes, no matter the size. $D_0 = \lim_{\epsilon \to 0} -\ln p / \ln \epsilon = 0$. If you wanted to work with a line of length l,

$$\begin{aligned} D_0 &= \lim_{\epsilon \to 0} -\frac{\ln l/\epsilon}{\ln 1/\epsilon} \\ &= \lim_{\epsilon \to 0} -\frac{\ln l + \ln 1/\epsilon}{\ln 1/\epsilon} \\ &= 1. \end{aligned}$$

Far more interesting is something like the Middle Third Cantor Set, where the lines do not cover the entire space. In fact, for the nth step, it takes 2^n lines to cover the length $3^{-n}l$. For this case

$$\begin{aligned} D_0 &= \lim_{n \to \infty} \frac{\ln 2^n}{\ln 1/3^{-n}} \\ &= \lim_{n \to \infty} \frac{n \ln 2}{n \ln 3} \\ &= \frac{\ln 2}{\ln 3} \\ &\approx 0.631. \end{aligned}$$

This result lies between the dimensions of a point and line ($0 < D_0 < 1$). Other structures can similarly exist between a surface and volume in dimension. Several features of this set are important to remember:

- the process is multiplicative,
- it is self-similar (very important),
- the result is patchy,
- this is a power law ($N(\epsilon) \sim \epsilon^{-D_0}$)
- and the effective dimensionality result need not be a whole number.

The simplest fractal model is the β model[6], using the active volume fraction $0 < \beta < 1$. For any dissipation eddy of size l in 3 dimensions, when it breaks into "daughter" eddies of size $l/2$, it must break into 8 in order to fill the same space. The fraction of space that is actually filled is then

$$\begin{aligned} \beta &= \frac{N}{2^3} \\ &= \frac{2^D}{2^3} \\ &= 2^{D-3}. \end{aligned}$$

Here we are using the fractal dimension D. If we look at an eddies "children" after n dissipative steps, we find $\beta_n = \beta^n = 2^{n(D-3)}$.

Looking at the Vorticity Equation for an incompressible fluid — obtained by taking the curl of the Navier Stokes Equation — we can see how vortex tube stretching relates to intermittency.

$$\frac{\mathrm{d}\vec{\omega}}{\mathrm{d}t} = (\vec{\omega} \cdot \vec{\nabla})\vec{v} + \nu^2 \nabla^2 \vec{\omega}$$

Because the vorticity is a derivative (curl) of the flow, we can look at

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} \sim \omega^2 + \nu^2 \nabla^2 \omega$$

as a vague approximation. This increases exceedingly fast, causing the magnitude to explode in a finite time. Here the mean dissipation rate is

$$\bar{\epsilon} \sim \beta_n \frac{v_n^3}{l_n},$$

but the occupation factor $2^{n(D-3)}$ can also be written as $(l_n/l_0)^{3-D}$. This is important, as it means there is a memory of the original size. To get the energy we rewrite the rate

$$\bar{\epsilon} \sim (\frac{l_n}{l_0})^{3-D}\frac{v_n^3}{l_n}$$

in terms of flow

$$v_n \sim (\bar{\epsilon} l_n)^{1/3}(\frac{l_n}{l_0})^{(D-3)/3}.$$

Then the energy is

$$\begin{aligned} E_n &\sim \beta_n v_n^2 \\ &\sim \bar{\epsilon}^{2/3} l_n^{2/3}(l_n/l_0)^{(3-D)/3}. \end{aligned}$$

$$E(k) \cong \bar{\epsilon}^{2/3} k^{-5/3}(kl_0)^{(D-3)/3}$$

The deviation between this result and the $-5/3$ from Kolmogorov is largest at the dissipation scale

$$\frac{\nu}{l_d^2} \sim \frac{v(l_d)}{l_d}$$

$$\rightarrow l_d \sim l_0 R_e^{-3/(1+D)}.$$

The higher order moments also deviate most at small scales:

$$\langle |\delta v_n|^p \rangle \cong \bar{\epsilon}^{-p/3} l_n^{p/3}(\frac{l_n}{l_0})^{\phi_p}.$$

Here we use $\phi_p \equiv (3-D)(3-P)/3$.

Throughout this example have been some important things to remember:

- intermittency is rooted in the active volume,
- there is an explicit memory and dependence on original scale l_0,
- cascades are self-similar, fractal structures ($D < 3$),
- this is a highly nonlinear, localized process,
- and the deviation from Kolmogorov's result is entirely dependent on the dimensionality.

Bifractals and Multifractals

What happens if there are multiple dissipative processes with different dimensions? Multifractals! For example, think of shocks: essentially the ramp is 1D and the drop is 0D. This type of bifractal system would need to have a phase transition or kink in it.

$$\frac{\delta v_l}{v_0} \sim \begin{cases} (l/l_0)^h 1 & \mathcal{L}_1, \, \mathrm{dim} D_1 \\ (l/l_0)^h 2 & \mathcal{L}_2, \, \mathrm{dim} D_2 \end{cases}$$

Now the union $\mathcal{L}_1 \cup \mathcal{L}_2$ must cover all of the dissipation. While the moments are a combination of terms

$$\frac{\langle \delta v_l^p \rangle}{v_0^p} = \mu_1 (\frac{l}{l_0})^{ph_1+3-D_1} + \mu_2 (\frac{l}{l_0})^{ph_2+3-D_2},$$

in reality, the one with the smallest power dominates.

$$\epsilon_p = min[ph_1 + 3 - D_1, ph_2 + 3 - D_2]$$

Where a kink in the structure function leads to a bifractal system, a curve or ensemble of kinks leads to a multifractal. Zonal flows, turbulence, and some waves in plasma can be thought of as multifractal systems. In these systems, you could (in theory) try to construct the spectrum of dimensionality by fitting higher order moments. On the other hand, this is a method that “fits everything, but explains nothing.”

References

[1] B. B. Mandelbrot, *Fractals and Scaling in Finance*, ch. States of randomness from mild to wild, and concentration from the short to the long run, pp. 117–145. New York, NY: Springer New York, 1997.

[2] “Central limit theorem.”

[3] S. Chandrasekhar, “Stochastic problems in physics and astronomy,” *Rev. Mod. Phys.*, vol. 15, pp. 1–89, Jan 1943.

[4] Y. B. Zel'dovich, S. A. Molchanov, A. A. Ruzmaĭkin, and D. D. Sokolov, “Intermittency in random media,” *Soviet Physics Uspekhi*, vol. 30, pp. 353–369, may 1987.

[5] A. N. Kolmogorov, "The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers," *Proceedings: Mathematical and Physical Sciences*, vol. 434, no. 1890, pp. 9–13, 1991.

[6] U. Frisch, P.-L. Sulem, and M. Nelkin, "A simple dynamical model of intermittent fully developed turbulence," *Journal of Fluid Mechanics*, vol. 87, no. 4, p. 719–736, 1978.