about everything; he is a pack of intuition; he is encyclopedic and is a universal conversationalist. If you manage to age well, you actually get better because you know so many more things. And he has an astonishing memory ("*une memoire d'éléphant*").

Having read descriptions of his personality, I was taken aback by the difference between the real man and the reputation of "arrogance" – which to me (as I am familiar with such accusations) comes merely from his targeted irreverence and lack of willingness to put up with established truths and established gods. People readily mistake irreverence towards some class of accepted heroes for arrogance. A fair approach would be to examine the targets of such irreverence.

In a way BM is the exact opposite of what I call the academic clerk: someone who is there to work on research like an obedient tax accountant. BM is a maverick, tenacious, and idiosyncratic in his approach; he seems to scorn formalities. It is all-natural that he would have had to counter resistance from the clerks. I was in for a surprise: I had the feeling of talking to a trader, capable of revising his views at a blip. And the man was simple, friendly, charming, the reverse of arrogant – except for his colorful irreverence. Consider that one of his colleagues, Michael Frame[8] who was also told that BM was "arrogant", accounts for his surprise upon having to contradict BM on a critical point. BM's reply was "Marvelous. The problem is more interesting than I had expected".

One final remark about recognition. When Daniel Kahneman received the Nobel medal many people congratulated him on such an honor. My reaction was to congratulate the Nobel committee: finally, these Swedes seem to be serious about their prize. Not only have they helped to make economics more of a science, but they also gave it the credentials to help enrich other disciplines. The abundance of data makes the field of economics an ideal laboratory to develop insights and quantitative tools helpful to other sciences – we can develop insights about human nature from economic choices (Kahneman and Tversky); we can also learn new mathematical methods (Mandelbrot). I hereby ask the Swedes to take some perspective and think of those whom, a century from now, will be identified as having changed the way we view the world. ■ *The (Mis)Behavior of Markets: A Fractal View of Risk, Ruin, and Reward* by Benoit Mandelbrot & Richard Hudson, Basic Books.

FOOTNOTES

1 Kenneth Falconer, *Nature*, 430/ 1 July 2004.
2 A shorter version of this book review was withdrawn from the *Los Angeles Times*, partly because I got too close to Mandelbrot after writing the review and did not want to bear the risk of personal conflict.
3 Roger Penrose, 2005, *The Road to Reality*, New York: Knopf.
4 John Brockman, 2005, Discussion with Benoit Mandelbrot, www.edge.org
5 See the posthumous *Probability Theory: The Logic of Science* by E.T. Jaynes, 2003, Cambridge University Press.
6 See Mandelbrot's essay on www.edge.com
7 *Journal of Economic Perspectives* Vol. 18, No. 3, Summer 2004
8 See the personal testimony in Michel L. Lapidus (editor): Fractal geometry and applications: A Jubilee of Benoit Mandelbrot, *Proceedings of Symposia in Pure Mathematics*, 72, 1, American Mathematical Society.

Fat Tails, Asymmetric Knowledge, and Decision Making

Nassim Nicholas Taleb's Essay in honor of Benoit Mandelbrot's 80th birthday

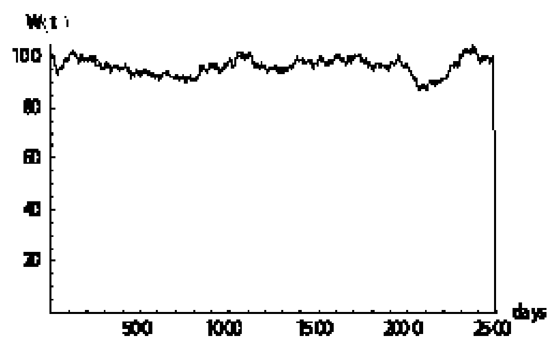

Figure 1: A dataset of 2,500 prices. Infer the attributes.

Introduction

Consider the following thought experiment. You show an agent a set of data of 2,500 days worth of returns (the resulting asset price $W(t)$ being represented in Figure 1) and ask him to infer the attributes of what he saw.

Odds are that he would tell you that the log-returns are Gaussian. 2,500 days data set represents an ample sample size by any measure, enough for the distribution to reveal itself to us. Clearly all the attributes of a mild distributions are there: no excess Kurtosis over that of a Normal, no outliers, no jumps, no gaps; a histogram of the returns would reveal the Platonic Bell Shape.

Now we continue with the rest of the story. We add one day, number 2,501; one single day can show a quite different picture.

Picture 2 shows the informational increase by that one day. The generating process for these draws is a mere switching process, built around a Gaussian, to which was added the occasional drawing, once in 2,500 days, from an infinite variance kick. This implies that the total is of infinite variance. Those who have not seen any such situation should take a look at emerging market currencies (those in a managed regime). It can also apply to a hedge fund returns: The properties of the late hedge fund LTCM are not too different from what we just saw. The bigger the divergence between the two regimes (the "normal" and the "unusual"), the worse the epistemological picture as more people will tend to be fooled by what they saw.

The central problem of uncertainty

What I call the central epistemological problem of uncertainty[1] is summarized as follows: we do not observe probability distributions, only random draws from an unspecified generator. So we need data to figure out the probability distribution. How do we gauge the sufficiency of the size of the sample? Well, from the probability distribution. If at the same time one needs data to figure out the probability distribution, and the probability distribution to figure out if we have enough data, then we have a severe circular epistemological problem.

Note here that fat tails are contagious. If you combine two random variables each following a power law distribution but with different exponents, the result is a power law distribution with, for tail exponent, the lower of the two. Here we have two processes, one of finite, the other of infinite variance; accordingly the infinite variance will prevail.

A traditional philosophical way to deal with the regress argument, if one follows the epistemological traditions, would be to either 1) put your hands up and bemoan the Problem of Induction, and find theological arguments to have some unquestioned belief or 2) proceed to a systematic layering: One can pose a meta-distribution, one that would take into account the probability of the candidate distribution being the wrong one. You can use priors and probabilize with series of meta-probabilities. Neither handy, nor convincing, and it implies as Elie Ayache[2] put it in this magazine "trying to find a random generator behind the random generator". And it does not escape the attacks by classical Pyrrhonian skeptics: we seem to be either 1) justifying belief with reference of other belief, itself justified by other belief, all the way up until some unargued dogma, which could be fragile (in this case some "known" distribution or generator for the time series) 2) justifying belief somewhere in the loop with another previously derived belief and falling back into severe circularity; finally 3) the regress may never end and we stay at the beginning.

Note that the quantitative-statistical literature is not thoughtful enough or self-critical "to be even wrong" on the subject. How? Conventional tests of normality study the square errors from a Gaussian and use a Gaussian-inspired distribution (a special case of the Gamma distribution, the Chi-Square, which is the distribution of the squared Gaussian variate) . This is exceedingly circular and reflects a severe lack of awareness of such circularity.

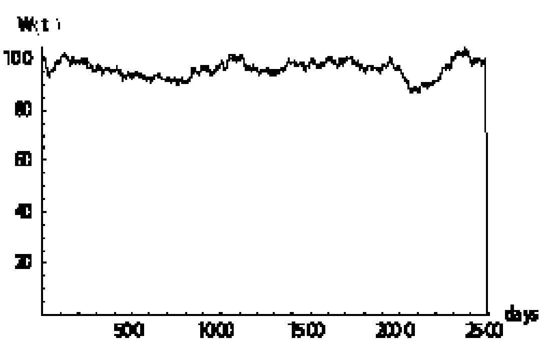

Figure 2: A dataset of 2,501 prices. What is the informational increase?

An easier solution

As an operator first and last, I believe that there are, however, far more elementary (and practical) ways to deal with this problem, or at least to protect ourselves from its ill effects. How? I propose two approaches.

First, consider Pascal's wager. We can change our payoff structure to accommodate what absence of knowledge we suffer from, and with respect to which moments of the distribution. For instance, if the data has "infinite" (or undefined) variance, one can avoid exposure to such infinite tail by clipping the sensitivity to the offending part of the distribution. Purchasing a simple derivative(say, an extremely out-of-the-money call), if it such product is available, may provide a solution. Our doubt can be targeted and remedied by transactions. *Tout simplement.*

Second, what we call the masquerade problem. The data cannot tell us what is the probability distribution generating it; but it can easily tell us what such probability distribution *is not* (or is not likely to be), and which moments of the distributions we may not be able to compute.

Portfolios, infinite variance, and epistemic opacity

What many academic philosophers do not realize is that the limits of some knowledge may be of small moment. I would rather use my energy in changing my payoff structure rather than getting into intractable issues and playing philosophaster. My colleague, another option trader and empirical philosopher Rabbi Anthony ("Tony") Glickman (also a Talmudic scholar), explains quite eloquently that being an option trader gives someone a philosophical approach along "long gamma" lines, or, more formally in the decision theory literature: along a mindset focused on the convexity of payoffs. One comment I make here about Tony is that his definition of philosopher is similar to mine (and Mandelbrot's): a philosopher is someone who specializes in ideas, not in other people's ideas – like stamp collecting. Professional philosophers can be like parasites. To Tony, like for me, *being long an option in the tail (or more generally "long convexity") eliminates the need to try to figure out what we don't know*[3]. Only an option trader could understand that – that's what I am trying to generalize to all decision making under uncertainty and convey to nontraders in my forthcoming *The Black Swan*.

It is key that we operators and decision makers are capable of insulating ourselves from nasty parts of the distribution. It is a fact that a portfolio constituted of securities that have infinite variance does not need to have infinite variance. How? If you are short a call spread with the position strike K, described as short a call struck at K, long another call at $K + y$, you are "short volatility", but you are not exposed to infinite variance. Your payoff is capped. Furthermore: *the properties of your strategy are not fragile to parametric assumptions or choice of model*.

Note here, in the earlier thought experiment, that the moments of the distribution are very precarious; the loss L (taken in Log returns) is so large that the moments are insensitive to the probability of the big loss π. Indeed the pair $\pi\, L$ (probability times the payoff) is so large that we may never care about the size of the probability. It is so obvious that we should work to control L – or, if we can't, to only enter transactions where such L can be controlled.

Now the question: what if we can't insulate ourselves from such distributions? The answer is "do something else", all the way to finding another profession. Risk managers frequently ask me what to do if the commonly accepted version of Value-at-Risk does not work. They still need to give their boss *some* number. My answer is: clip the tails if you can; get another job if you can't. "Otherwise you are defining yourself as a slave". If your boss is foolish enough to want you to guess a number (patently random), go work for a shop that eliminates the exposure to its tails and does not get into portfolios first then look for measurement after. Indeed if like me you think that Modern Portfolio Theory is charlatanism (as confirmed by my trader's observations and empirical research, and Mandelbrot's work), use portfolios that do not depend on their measurements. It is so easy to avoid traps.

The asymmetric masquerade problem

A power law (as we saw in the thought experiment) can easily masquerade as a Gaussian but not the reverse (at least not easily). We can reject the Gaussian more easily than we can accept it. More generally, a distribution with fat tails can show milder tails than its "true" properties, except, of course, when it is too late. It will even tend to do so. The small sample properties of these processes are such that we are not likely to encounter large moves in them. We can call that problem an "epistemic headwind". To answers some questions put to me in this magazine about skepticism and asymmetric knowledge[4], I will use the argument that it is always easier to figure out what the distribution is not than what it is. Compare that to the attributes of humans: a criminal can masquerade as an honest citizen; an honest citizen cannot as easily fake being a criminal. Many extensions of this point are accepted in many fields: one single event constitutes a catastrophe; one needs many days without an event to pronounce an environment as catastrophe-free.

This asymmetry is at the core of skeptical empiricism: our body of knowledge is more readily increased by negative observations than by confirming ones. Remarkably, we can do something with this; it leads us to a ranking of the robustness of results. And remarkably, it is because I elect to behave operationally as if the market followed a Mandelbrot-stable process that I can build portfolios that I am comfortable with. A Mandelbrot-stable variable is simply here what is called a Levy stable, but with non-serially independent draws (what BM calls multifractal). We will return to the situation.

The α problem

Take X a random variable, we have a power law $P[X > x_0] \sim O(x_0^{-\alpha})$. Clearly we are told that if the first and second moments of the distribution are defined, i.e., $\alpha > 2$, then, under aggregation the series becomes Gaussian so we can use the conventional tools of analysis. Note here that this only holds if we have independent increments.

BM came up with papers in the 1960s[5] showing cotton prices with tail $\alpha < 2$, in other words implying Levy-stability; the distribution has fat tails and does not become a Gaussian under aggregation. There have been series of papers[6] disagreeing with Mandelbrot's early work and its conclusions. Researchers tend to be "skeptical" about the Lévy regime hypothesis producing, for more than a quarter century now, "evidence" to the effect that Mandelbrot's early characterization of infinite variance is wrong – people seem to very badly need a Gaussian in order for them to operate with the current academic framework. Their methodology is based on two arguments, first, the "observation" of $\alpha > 2$ and, second, the examination of the behavior of the data when they lengthen the time observation period.

These studies are either inconsequential or wrong in their inferences. First, it does not make much difference whether or not we are in a Lévy regime since we don't really stay in the Gaussian regime in the parts of the distribution that matter. Second, we do not "have evidence" that we are not in a Lévy regime. Third, we need to go beyond the "Lévy regime" and consider the Mandelbrot regime by lifting the too-restrictive assumption of independent increments. I will get into the details of the arguments next.

Point 1: The slowness in the rate of convergence makes a cubic α very seriously NonGaussian.
If we accept that α is approximately 3, "outside the Lévy regime", we are still in trouble with respect to the convergence to the Gaussian. Finite second moment implies convergence under aggregation, but we need to remember that with $\alpha < 4$ have an undefined 4th moment. The implication is rather serious. Consider that the 4th moment is the variance, corresponds to the error of the measurement in the variance (what we option traders call the "Vvol"). It will be infinite! This implies a quite nasty rate of convergence. There will always be a NonGaussian jump in the extreme tail to make the tail scalable.

Another way to view it is that the observations that we are adding are likely to be biased towards the middle of the distribution, making it converge in the body but much more slowly in the tails. We can examine this quantitatively. Take $\alpha = 3$. It is easy to show[7] that, in standard deviation terms, outside (Log (n), with n the number of observations, we stay in a scalable regime. Even if you add up 1 million days, the Gaussian

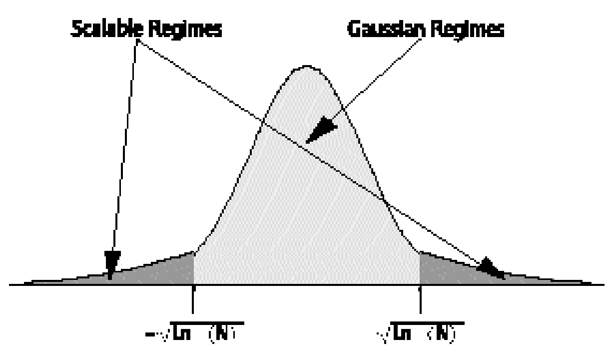

Figure 3: The regime densities.

regime stops at 3.7 sigmas! Typical penny-wisdom since the consequences of outside such moves are disproportionately large. Figure 3 shows the two-regime densities.

The situation is reminiscent of the value at risk problem. The tail of the distribution is where our errors compound. That is where ironically people like the precision.

Point 2: Absence of Evidence is Not Evidence of Absence: The Small Sample Bias Problem

Measuring an $\alpha > 2$ does not imply with any confidence that the "true" α is not <2. I avoid the discrepancies here in the measurement results from the various estimators, whether Hill and Log-Log linear regression. It just takes time (and data) for these distributions to reveal themselves. Simulate a series of symmetric random draws with $\alpha = 1.9$ and you will recover an α close to 3 with 106 samples. This is an argument well known to many traders[8] and discussed in Weron (2003).[9]

As we saw with the 2,500 day properties in the thought experiment, matters can be even more complex with a mixed process. In short, the fat tailed process tend to show the underestimation of the observed volatility.

Point 3: Where the Aggregation Fattens the Tails. Many of these inferences and indeed much of the mathematics we are used to assumes that we have independent draws

Now consider the following intuition: very bad moves generate very large up or down moves. And also consider that this may only happen in extreme circumstances, when the moves exceed a given threshold. Intuitively, a large loss might generate series of self-causing liquidations. What would that do to the scalability?

Well, in such mechanism, the aggregation *fattens* the tails. Such is the observation made by Sornette concerning the events leading up to the crash of 1987,[10] prompting him to analyze the properties of "drawdowns" independently.

This brings us to the Mandelbrot multifractal generalization[11] that shows that the process can have $1 < \alpha < \infty$. Indeed much of the work on stable distributions is restrictive – obsessively relying on the assumption of independence.

Final note and consequences for Financial Engineering and Quantitative Finance.

I conclude by saying that to many of us the field of finance seem to be intricately linked to modern portfolio theory. I showed that it does not have to be so. And it does not take much to fix the problem.

FOOTNOTES

1 See Taleb and Pilpel, 2004. See also Elie Ayache , 2004a
2 Ayache, 2004b.
3 We can extend this convexity argument to the philosophy of probability in general. Take the subjectivist concept of probability as degree of belief attributed to De Finetti. Probability is held to be the price I am willing to fix in such a way as I would equivalently buy or sell a state of the world, making sure I remain consistent and avoid the "Dutch book" problem. Well, you do not have to put yourself in a situation which you have to trade.
4 see Ayache 2004.
5 Mandelbrot (1963) See also Mandelbrot (1997).
6 See Officer (1972), Stanley et al (2000), Gabaix et al (2003)
7 See Sornette (2004) for the proof. See also Bouchaud and Potters (2003).
8 Mark Spitznagel brought this to my attention.
9 Weron (2001).
10 See Sornette (2004) for the argument.
11 Mandelbrot (2000a, 2000b).

REFERENCES

■ E. Ayache,2004a, The Back of Beyond, *Wilmott*, 26-29
■ E. Ayache,2004b, A Beginning, in the end, *Wilmott*, 6-11
■ M. Blyth, R. Abdelal and Cr. Parsons,2005, Constructivist Political Economy, Preprint, forthcoming, 2006: *Oxford University Press*.
■ J.-P. Boucheaud and M. Potters, 2003, *Theory of Financial Risks and Derivatives Pricing, From Statistical Physics to Risk Management*, 2 nd Ed., Cambridge University Press.
■ X. Gabaix, P. Gopikrishnan, V.Plerou & H.E. Stanley, 2003, A theory of power-law distributions in financial market fluctuations, *Nature*, 423, 267-270.
■ B. Mandelbrot, 1963, The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419.
■ B. Mandelbrot, 1997, *Fractals and Scaling in Finance*, Springer-Verlag.
■ B. Mandelbrot, 2001a, *Quantitative Finance*, 1, 113-123
■ B. Mandelbrot, 2001b, *Quantitative Finance*, 1, 124-130
■ R. R. Officer 1972 J. Am. Stat. Assoc. 67, 807–12
■ D. Sornette, 2003, *Why Stock Markets Crash : Critical Events in Complex Financial Systems*, Princeton University Press
■ D. Sornette ,2004, 2 nd Ed., *Critical Phenomena in Natural Sciences, Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*, (Springer Series in Synergetics, Heidelberg)
■ H.E. Stanley, L.A.N. Amaral, P. Gopikrishnan, and V. Plerou, 2000, Scale invariance and universality of economic fluctuations, *Physica A*, 283,31-41
■ N N Taleb and A Pilpel, 2004, I problemi epistemologici del risk management in: Daniele Pace (a cura di) Economia del rischio. Antologia di scritti su rischio e decisione economica, Giuffrè, Milano.
■ R. Weron, 2001,Levy-stable distributions revisited: tail index > 2 does not exclude the Levy-stable regime.*International Journal of Modern Physics C* (2001) 12(2), 209-223.