Lecture 3 The fundamental problem of the 0^{th} moment and the irrelevance of "naked probability"

The Nonbinary problem: Decisions (by humans) are rarely made based on probability except in the case of strictly binary bets: those on win/lose, in which the agent is focused on the outcome of naked probability $\int_D p(x)\,dx$ rather than $\int_D f(x)\,p(x)\,dx$ where f(x) is a function of the random variable and *D* the area of integration. So for the expectation, i.e. "impact", most common criterion of concern, f(x)= x.

The agent might discuss the "probable" and "improbable" but not know that he does not really mean it. It is just a proxy for something else -"consequential" or "inconsequential".

Note:We will see (section x) that we do not care about some part of the 0^{th} moment *(probability),* burt some part of the first moment, and **just that** --or some complicated function of it, causing the dependence on the $\mathcal{L}^1$ norm. We have no reasons (except computational) to worry about $\mathcal{L}^2$ or higher ones (higher moments). Anyway, we will be using f for the expectation of scaling of the outcomes. $f(x) = |x|$ (mean deviation) or $f(x)=x^2$ for the variance, etc., for higher moments, One can even include some "utility" function as part of *f*, whatever that means --it is not necessary as it can be embedded in p(x).

One aspect of this irrelevance of probability is that "fatter tails" does not necessarily mean a higher **incidence** *(i.e. frequency) of rare events; it means a higher* **contribution** *of these events which generally corresponds to a lower incidence of* **some** *tail events (and a rise of others further out in the distribution). Given the same scaling, a higher fourth moment "fatter tails" decreases the probability of exceeding K, i.e.,* $\int_K^\infty f(x)\,dx$ *while increasing the contribution* $\int_K^\infty x\, f(x)\,dx$

Example: Naive Fattening of the Gaussian

Create a naive fat-tailed Gaussian. We pick a dual Gaussian mixture, both mixes equiprobable ($\frac{1}{2}$) with a "low" variance $(\sigma(1-v))^2$ and a "high" one $\left(\sigma\sqrt{-v^2+2v+1}\right)^2$ selecting a single v so that the total variance remains the same. With 1 >v≥ 0, the total standard deviation

$$\sigma = \sqrt{\frac{1}{2}\left((\sigma(1-v))^2 + \left(\sigma\sqrt{-v^2+2v+1}\right)^2\right)}$$

v=0, v= 1/2 fattens the tails up to 1 standard deviation

Illustration: Time spent in the "tunnel" between -1 and 1 "sigmas" for the deterministic and mild Gaussian mixture

We can see that as v increases (therefore volatility is more stochastic), the time spend betwen +1 and -1 standard deviations increases. So events , like P>1σ , with 16% probability have actually 12% of occuring.

	v	Time ± 1 std
1	0	0.682689
2	0.1	0.687089
3	0.2	0.698764
4	0.3	0.715553
5	0.4	0.73477
6	0.5	0.752404
7	0.6	0.763293

Beyond some "sigma" the effect reverses -- here rather quickly: 3 standard deviations. So fatter tails imply fewer 1 sigma events, and more 3 sigma ones. Simply, we are not dealing with very fat tails as these do not fill out too far outside the central region.

	v	Time ± 3 MAD
1	0	0.983319
2	0.1	0.98198
3	0.2	0.978556
4	0.3	0.973975
5	0.4	0.969164
6	0.5	0.964807
7	0.6	0.961187

Stopping Time & Fattening of the tails of a Brownian Motion

Consider the distribution of the time it takes for a continuously monitored Brownian motion S to exit from a "tunnel" with a lower bound L and an upper bound H. Counterintuitively, fatter tails makes an exit (at some sigma) take longer. You are likely to spend more time inside the tunnel --since exits are far more dramatic.

ψ is the distribution of exit time t, where $t \equiv inf\{t: S \notin [L,H]\}$

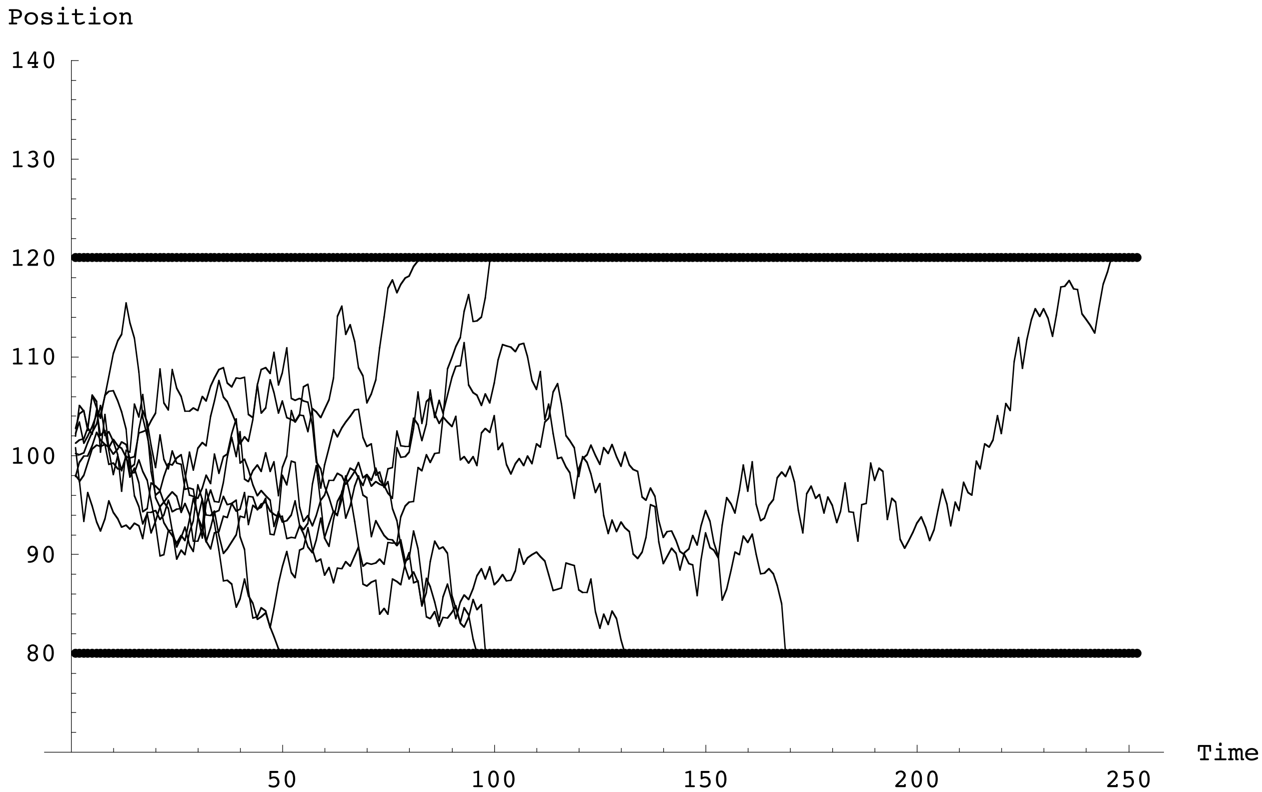
From Taleb (1997) we have the following approximation

$$\psi(t \mid \sigma) = \frac{1}{(\log(H) - \log(L))^2} \left(e^{-\frac{1}{8}(t\sigma^2)} \pi \sigma^2 \sum_{n=1}^{m} \frac{(-1)^n e^{-\frac{n^2 \pi^2 t \sigma^2}{2(\log(H)-\log(L))^2}} n \sqrt{S} \left(\sqrt{L} \sin\left(\frac{n\pi(\log(L)-\log(S))}{\log(H)-\log(L)}\right) - \sqrt{H} \sin\left(\frac{n\pi(\log(H)-\log(S))}{\log(H)-\log(L)}\right)\right)}{\sqrt{H}\sqrt{L}} \right)$$

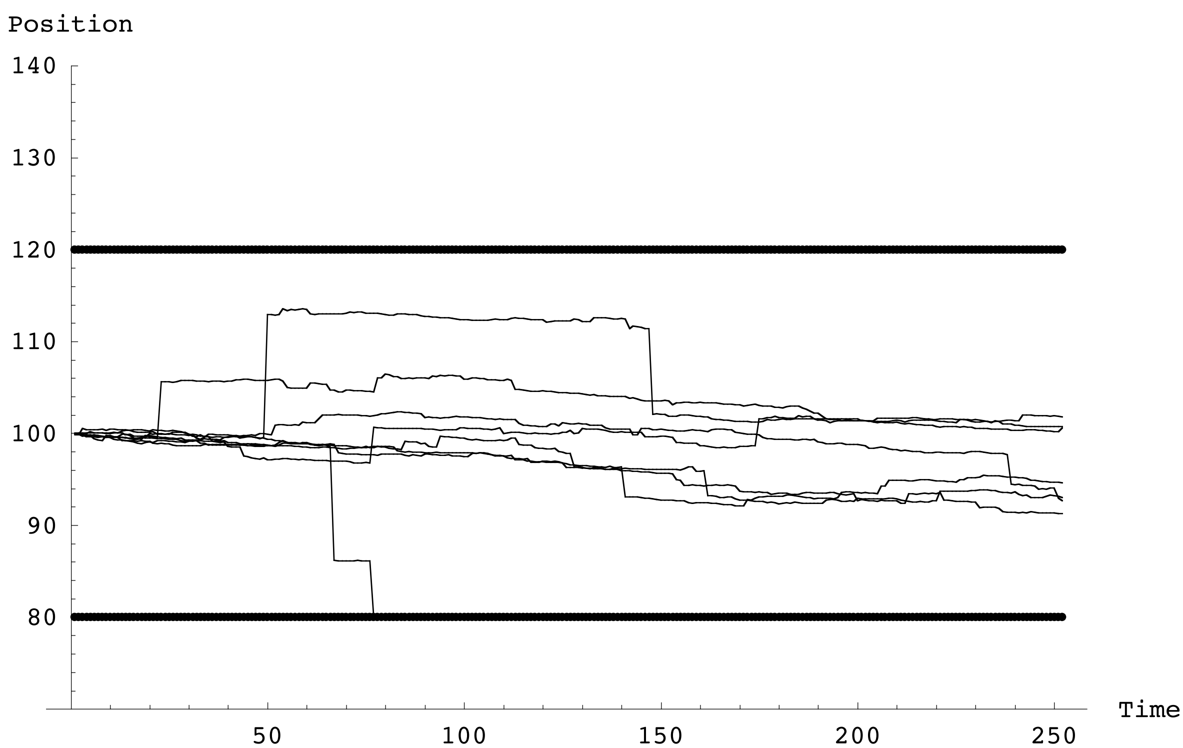
and the fatter-tailed distribution from mixing Brownians with σ^2 separared by a coefficient v:

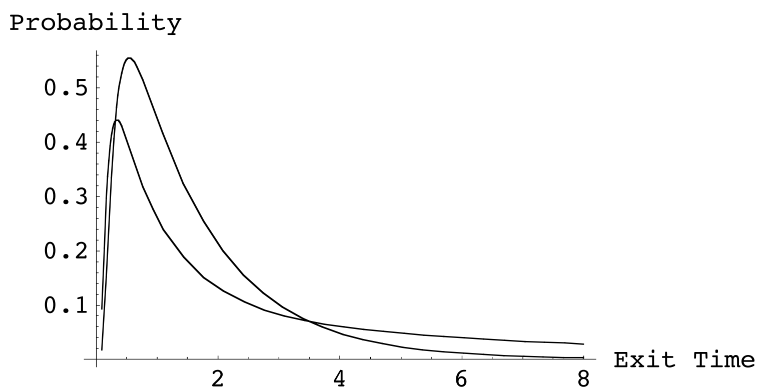
$$\psi(t \mid \sigma, v) = \frac{1}{2} p(t \mid \sigma(1-v)) + \frac{1}{2} p\left(t \,\middle|\, \sigma\sqrt{-v^2 + 2v + 1}\right)$$

Stochastic paths terminating upon hitting barriers H (high) H=120 and L (low) L=80. Time to exit is extended by the fattening of the tails.

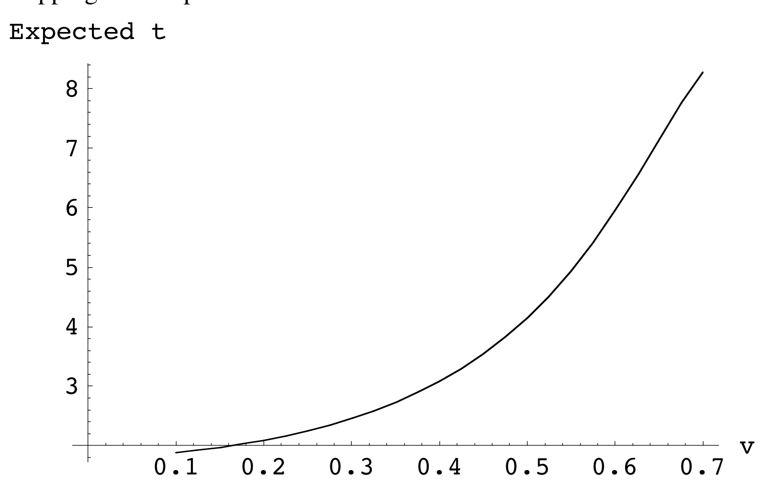


With very fat tails (almost-Cauchy)





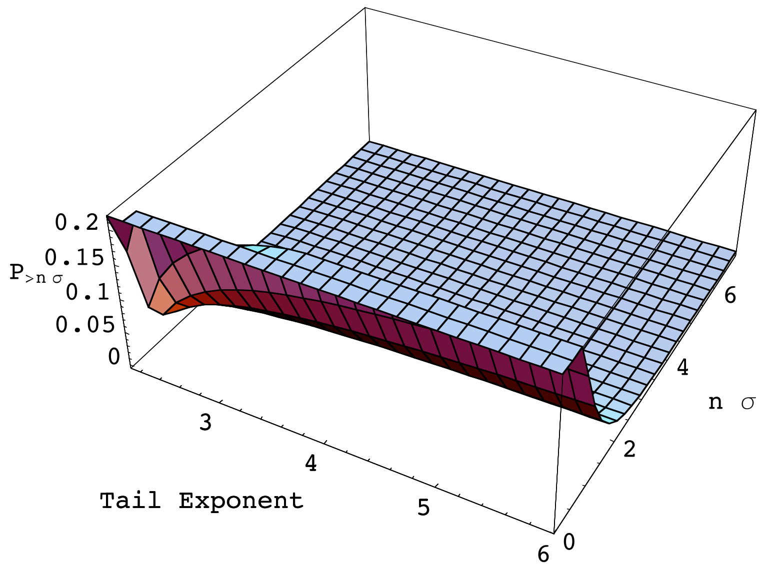
Expected stopping time explosion with v



The Implication: Visibly it "takes longer" to capture the statistical properties. How much "longer"? I don't know --it is, as we saw earlier, an inverse problem.

Fatter tails, or "truer" fat tails

Here I use a finite variance scalable distribution. Clearly at some exponent, the probability of exceeding drops for for n σ low and rises for higher n σ



Discrete Calc