“Prediction Markets” & Game Setups Have Nothing to Do With Real Risks

A SHORT EXPLANATORY TECHNICAL NOTE, NOT A PAPER

Nassim N. Taleb
NYU-Poly

November 2012

Abstract

This explains how and where prediction markets (or, more general discussions of betting matters) do not correspond to reality and have little to do with exposures to fat tails and “Black Swan” effects. Elementary facts, but with implications. This show show, for instance, the “long shot bias” is misapplied in real life variables, why political predictions are more robut than economic ones.

This discussion is based on Taleb (1997) showing the difference between a binary and a vanilla option.

Definitions

A binary bet (or just “a binary” or “a digital”): a outcome with payoff 0 or 1 (or, yes/no, -1,1, etc.) Example: “prediction market”, “election”, most games and “lottery tickets”. Also called **digital**. Any statistic based on YES/NO switch.

Binaries are effectively bets on probability. They are rarely ecological, except for political predictions.

(More technically, they are mapped by the Heaviside function.)

A exposure or “vanilla”: an outcome with no open limit: say “revenues”, “market crash”, “casualties from war”, “success”, “growth”, “inflation”, “epidemics”... in other words, about everything.

Exposures are generally “expectations”, or the arithmetic mean, never bets on probability, rather the pair probability × payoff

A bounded exposure: an exposure(vanilla) with an upper and lower bound: say an insurance policy with a cap, or a lottery ticket. When the boundary is close, it approaches a binary bet in properties. When the boundary is remote (and unknown), it can be treated like a pure exposure. The idea of “clipping tails” of exposures transforms them into such a category.

The Problem

The properties of binaries diverge from those of vanilla exposures. This note is to show how conflation of the two takes place: prediction markets, ludic fallacy (using the world of games to apply to real life),

1. They have diametrically opposite responses to skewness.
2. They repond differently to fat-tailedness (sometimes in opposite directions). Fat tails makes binaries more tractable.
3. Rise in complexity lowers the value of the binary and increases that of the exposure.

Some direct applications:

1- Studies of “long shot biases” that typically apply to binaries should not port to vanillas.

2- Many are surprised that I find many econometricians total charlatans, while Nate Silver to be immune to my problem. This explains why.

3- Why prediction markets provide very limited information outside specific domains.

4- Etc.

The Elementary Betting Mistake

One can hold beliefs that a variable can go lower yet bet that it is going higher. Simply, the digital and the vanilla diverge. $P(X > X_0) > \frac{1}{2}$, but $E(X) < E(X_0)$. This is normal in the presence of skewness and extremely common with economic variables. Philosophers have a related problem called the lottery paradox which in statistical terms is not a paradox.

The Elementary Fat Tails Mistake

A slightly more difficult problem. When I ask economists or social scientists, “what happens to the probability of a deviation >1σ when you fatten the tail (while preserving other properties)?”, almost all answer: it increases (so far all have made the mistake). Wrong. They miss the idea that fat tails is the contribution of the extreme events to the total properties, and that it is a pair probability × payoff that matters, not just probability.

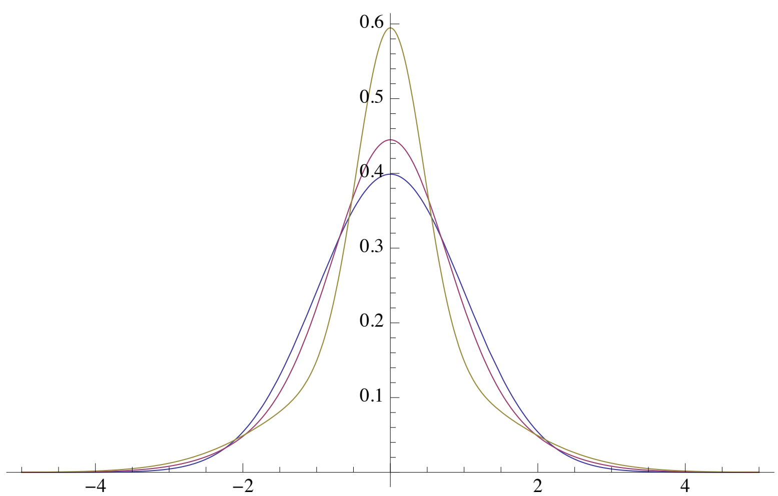
I’ve asked variants of the same question. “The Gaussian distribution spends 68.2% of the time between ± 1 standard deviation. The real world has fat tails. In finance, how much time do stocks spend between ± 1 standard deviations?” The answer has been invariably “lower”. Why? “Because there are more deviations.” Sorry, there are fewer deviations: stocks spend between 78% and 98% between ± 1 standard deviations (computed from past samples).

Some simple derivations: Let x follow a Gaussian distribution (μ , σ). Assume μ=0 for the exercise. What is the probability of exceeding one standard deviation? $P_{>1\sigma} = 1 - \frac{1}{2}\,\mathrm{erfc}\left(-\frac{1}{\sqrt{2}}\right)$, where erfc is the complimentary error function, $P_{>1\sigma} = P_{<1\sigma} \simeq 15.86\%$ and the probability of staying within the “stability tunnel” between ± 1 σ is ≃ 68.2 %.

Let us fatten the tail, using a standard method of linear combination of two Gaussians with two standard deviations separated by $\sigma\sqrt{1+a}$ and $\sigma\sqrt{1-a}$, where a is the "vvol" (which is variance preserving, a technically of no big effect here, as a standard deviation-preserving spreading gives the same qualitative result). Such a method leads to immediate raising of the Kurtosis by a factor of $(1+a^2)$ since $\frac{E(x^4)}{E(x^2)^2} = 3\,(a^2+1)$

$$P_{>1\sigma} = P_{<1\sigma} = 1 - \frac{1}{2}\,\mathrm{erfc}\left(-\frac{1}{\sqrt{2}\,\sqrt{1-a}}\right) - \frac{1}{2}\,\mathrm{erfc}\left(-\frac{1}{\sqrt{2}\,\sqrt{a+1}}\right)$$

So then, for different values of a as we can see, the probability of staying inside 1 sigma increases.



Fatter and fatter tails: different values of *a*. We notice that higher peak ⟹lower probability of nothing leaving the ±1 σ tunnel

The Event Timing Mistake

Fatter tails increases time spent between deviations.

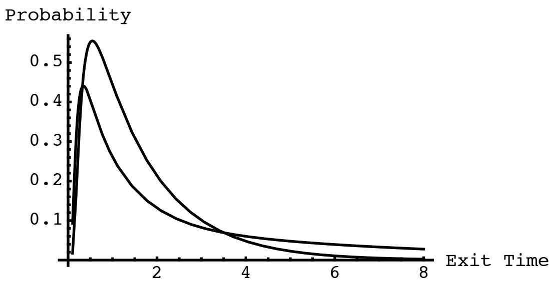
Stopping Time & Fattening of the tails of a Brownian Motion: Consider the distribution of the time it takes for a continuously monitored Brownian motion S to exit from a "tunnel" with a lower bound L and an upper bound H. Counterintuitively, fatter tails makes an exit (at some sigma) take longer. You 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

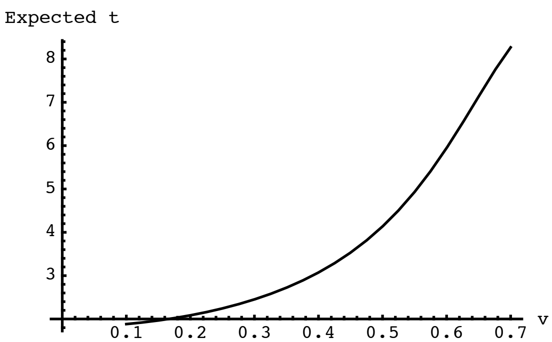
$$\psi(t \mid \sigma) = \frac{1}{(\log(H)-\log(L))^2}\left(e^{-\frac{1}{8}(t\sigma^2)}\,\pi\,\sigma^2 \sum_{n=1}^{m} \frac{1}{\sqrt{H}\,\sqrt{L}}\,(-1)^n\, e^{-\frac{n^2\pi^2 t\sigma^2}{2(\log(H)-\log(L))^2}}\, n\sqrt{S}\left(\sqrt{L}\,\sin\left(\frac{n\pi(\log(L)-\log(S))}{\log(H)-\log(L)}\right) -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

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

$$\psi(t \mid \sigma, a) = \frac{1}{2}\, p(\, t \mid \sigma\,(1 - a)) + \frac{1}{2}\, p(t \mid \sigma\,(1 + a))$$

This graph shows the lengthening of the stopping time between events coming from fatter tails.





{this is a note. A more advanced paper explains why more uncertain mean (vanilla) might mean less uncertain probability (prediction), etc. Also see the “Why We Don’t Know What We Are Talking About When We Talk About Probability”.