Fluctuations of Martingales and Winning Probabilities of Game Contestants

David Aldous and Mykhaylo Shkolnikov

August 1, 2013

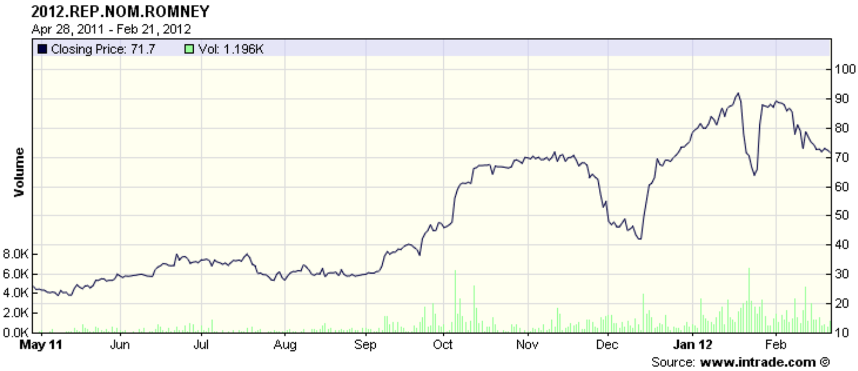
An interesting topic for exposition – can discuss in

- Popular talk (to non-mathematicians)
- Undergraduate stochastic processes course
- Graduate continuous martingale course
- Research problems we can do
- Open research problems

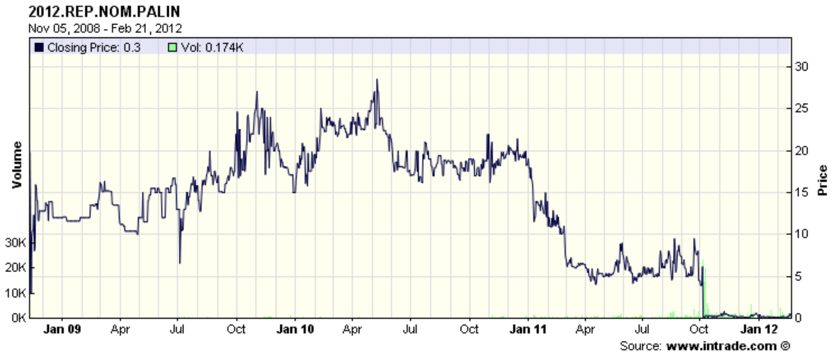
Based on two papers

- Using prediction market data to illustrate undergraduate probability. Amer. Math. Monthly (2013), to appear.
- Fluctuations of martingales and winning probabilities of game contestants. Electronic J. Probability (2013)

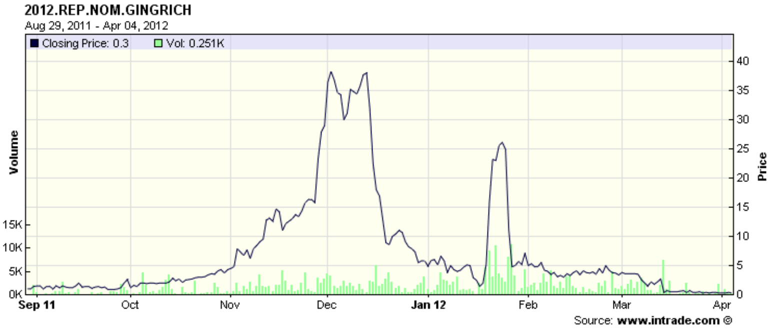
Romney for 2012 Republican presidential nominee.
Intrade prediction market prices through end February 2012.



Palin for Republican presidential nominee.



Gingrich for Republican presidential nominee.



Widely believed that (in this particular race) the number of candidates whose prospects rose and then and fell was unusually large. But was it really?

A prediction market price of (say) 57 reflects a consensus estimate of 57% probability. I talk about “prices” to avoid talking about “probabilities of probabilities”.

The **conservation of fairness principle** (optional sampling theorem) easily gives a theoretical formula:

if each candidate's initial price is below x then the mean number of candidates whose price ever reaches or exceeds x equals $100/x$.

Data: Maximum price for each candidate.

Romney 100
Perry 39
Gingrich 38
Palin 28
Pawlenty 25
Santorum 18
Huntsman 18
Bachmann 18
Huckabee 17
Daniels 14
Christie 10
Giuliani 10
Bush 9
Cain 9
Trump 8.7
Paul 8.5

Conclusion: Only slightly more than expected.

We can conduct another check of theory versus data by considering downcrossings. The hypothesis implies:

> *Consider a price interval $0 < a < b < 100$, and consider an upcoming election with several candidates, and a (prediction market) price for each candidate, where initially all these prices are below b. Theory says that the expected total number of downcrossings of prices (sum the numbers for each candidate) over the interval $[a, b]$ equals $(100 - b)/(b - a)$.*

To gather data for the interval $[10, 20]$, we need only look at the five candidates whose maximum price exceeded 20, and their numbers of downcrossings of $[10, 20]$ were:

Palin (2); Romney (0); Perry (1); Pawlenty (2); Gingrich (2).

So the observed total 7 is in fact close to the theoretical expectation of 8.

The math setup below models contestants in a contest which will have one winner at some unknown random future time – $M_i(t)$ is the probability that contestant i will be the winner, given information known at time t. In this scenario all the assumptions will hold automatically except for path-continuity, which expresses the idea that information becomes known slowly.

Math setup. Given a probability distribution $\mathbf{p} = (p_i, i \geq 1)$ consider a collection of processes $(M_i(t), 0 \leq t < \infty, i \geq 1)$ adapted to a filtration $(\mathcal{F}_t)$ and satisfying
(i) $M_i(0) = p_i, i \geq 1$;
(ii) for each $t > 0$ we have $0 \leq M_i(t) \leq 1 \ \forall i$ and $\sum_i M_i(t) = 1$;
(iii) for each $i \geq 1$, $(M_i(t), t \geq 0)$ is a continuous path martingale;
(iv) there exists a random time $T < \infty$ a.s. such that, for some random I, $M_I(T) = 1$ and $M_j(T) = 0 \ \forall j \neq I$.

Call such a collection a $\mathbf{p}$-*feasible* process, and call the $M_i(\cdot)$ its component martingales.

For $0 < a < b < 1$ consider

$$N_b := \text{number of } i \text{ such that } \sup_t M_i(t) \geq b$$

$$D_{a,b} := \text{sum over } i \text{ of the number of downcrossings of } M_i(\cdot) \text{ over } [a,b].$$

Here's what we said before.

Lemma

If $\max_i p_i \leq b$ then for any $\mathbf{p}$-feasible process,

$$\mathbb{E}N_b = 1/b, \quad \mathbb{E}D_{a,b} = (1-b)/(b-a).$$

In contrast, the distributions of N_b and $D_{a,b}$ will depend on the joint distributions of the component martingales.

Question: what are the extremal possibilities, measured (say) by variance?

Imagine a multi-episode TV show which starts with M contestants and ends with one winner. Suppose the contestants are equally good at whatever games/challenges are used. Mimic by some "purely random" scheme such as:

1: Survivor scheme. Each week, pick random one remaining contestant and eliminate that person.

2: . . . . . . Millionaire scheme. (variant where forced to try for the million; season ends when won). Each week, pick random contestant; that person either wins or is eliminated, fairly.

(Of course these are discrete – an exercise (graduate course) to embed into our continuous setting.)

In the Survivor scheme there are times

$$\ldots \leq T_4 \leq T_3 \leq T_2 \leq T_1$$

where at time T_k there are k remaining contestants, each with "price" $1/k$. So, regardless of how we embed, the only possible values of N_b are $\lfloor 1/b \rfloor$ and $\lceil 1/b \rceil$, and we know $\mathbb{E}N_b = 1/b$, so this attains the minimum possible variance.

What is N_b in the Millionaire scheme?

If there are $k > 1/b$ contestants remaining, then for the current contestant there are 3 possibilities:

reach price b, go on to win: chance $\frac{1}{kb} \times b$
reach price b, go on to be eliminated: chance $\frac{1}{kb} \times (1 - b)$
eliminated without reaching price b: chance ......

We recognize this as (essentially) the **craps principle** setting for the Geometric(b) distribution; and in fact the limit distribution of N_b as the initial number contestants $\to \infty$ is exactly Geometric(b).

A half-page argument shows this is the worst case:

Proposition

Any possible limit distribution for feasible processes has variance at most $(1 - b)/b^2$, the variance of Geometric(b).

This indicates that, within the class of **p**-feasible processes there are two special processes, Survivor and Millionaire, which are in certain senses (variance of N_b) at opposite extremes within the class. But is there some special "canonical" process within the class?

A textbook example of a martingale is the 2-allele Wright-Fisher chain with no mutation or selection; the infinite-population limit is the Wright-Fisher diffusion on $[0, 1]$. There is an analogous n-allele version, initial frequencies $\mathbf{p} = (p_1, \ldots, p_n)$, and the diffusion limit is a **p**-feasible process with nice mathematical properties.

Open problem. Calculating distribution of N_b for the Wright-Fisher process in the limit $\max_i p_i \to 0$ involves solving a PDE . . . . . .

Recall that the expectation $\mathbb{E}D_{a,b}$

$$D_{a,b} = \text{ sum over } i \text{ of the number of downcrossings of } M_i(\cdot) \text{ over } [a,b]$$

does not depend on the **p**-feasible process. Aldous - Shkolnikov (2013) give upper and lower bounds on the variance of $D_{a,b}$; the Survivor and Millionaire process variances are fairly close to the bounds, but we do not know what the precise extremal processes are.

To see why this is not elementary, if we consider the "infinitesimal interval" limit as $a \to b$, and then take b small, then we are led to the following "cleaner" question. Recall: for standard Brownian motion, run until hitting -1,

$$L := \text{total local time at } 0$$

has Exponential distribution.

Now consider k processes, each standard Brownian motion w.r.t. the same filtration, but otherwise arbitrarily dependent. Consider the local times $L_1, \ldots, L_k$ for each process.

Open Problem. What is the minimum possible value of $\mathrm{var}(\sum_{i=1}^k L_i)$?

Math setup – repeat previous slide.

Given a probability distribution $\mathbf{p} = (p_i, i \geq 1)$ consider a collection of processes $(M_i(t), 0 \leq t < \infty, i \geq 1)$ adapted to a filtration $(\mathcal{F}_t)$ and satisfying

(i) $M_i(0) = p_i, i \geq 1$;

(ii) for each $t > 0$ we have $0 \leq M_i(t) \leq 1 \; \forall i$ and $\sum_i M_i(t) = 1$;

(iii) for each $i \geq 1$, $(M_i(t), t \geq 0)$ is a continuous path martingale;

(iv) there exists a random time $T < \infty$ a.s. such that, for some random I, $M_I(T) = 1$ and $M_j(T) = 0 \; \forall j \neq I$.

Call such a collection a $\mathbf{p}$-*feasible* process, and call the $M_i(\cdot)$ its component martingales.

Some more abstract theory

In our math setup we were given the initial values (p_i) for the probability that contestant i is the winner. Suppose (p_i) is uniform on n contestants. Can we let $n \to \infty$ and consider starting with an infinite number of contestants, each with the same infinitesimal chance of winning?

(A) For the two particular processes (Survivor/Millionaire), easy to do an inductive construction of such a process.

(B) Recall the n-allele Wright-Fisher diffusion. Ethier-Kurtz (1981) showed the infinite-allele analog exists (represented by ranked frequencies). Our context is artificial as genetics but analogous results arise in math of "stochastic coalescence". **But** such results are all in the Markovian setting – determining the entrance boundary of a given Markov process.

(C) Aldous-Shkolnikov (2013) formalizes the notion of a **0**-feasible process $(M_i(t), 0 < t < \infty)$ which on $t_0 \leq t < \infty$ behaves as some **p**-feasible process but which has $\max_i M_i(t) \to 0$ a.s. as $t \downarrow 0$. This is more subtle than it looks!