Thresholds in Bak–Sneppen type models

M. Grinfeld and P. Knight

Mathematics, University of Strathclyde

Plan

1. Apology
2. Apology
3. Bak-Sneppen type models
4. Construction
5. Markov chain definition
6. Approximating the threshold: an “ergodicity” condition
7. A coagulation-fragmentation framework
8. Remarks

Apologies 1 and 2

- MG, L. Piscitelli, and R. Cross, Physica A **287** (2000), 577-586.
- R. Cross, MG, H. Lamba, and A. Pittock, Rationality, frustration minimisation, hysteresis and the El Farol problem, in *Relaxation Oscillations and Hysteresis*, SIAM, Philadelphia 2005, pp. 61–72.
- R. Cross, MG, H. Lamba, and T. Seaman, Physica A **354** (2005), 463–478.
- R. Cross, MG, and H. Lamba, J. Phys.: Conference Series **55** (2006), 55–62.
- R. Cross, M. Grinfeld, H. Lamba, and T. Seaman, Euro. Phys. J. B **57** (2007), 213–218.
- R. Cross, MG, and H. Lamba, IEEE CSM **29** (2009), 30–43.

This is a financial markets modelling framework in the shape of a mean-field stochastic HAM.

Good things: Very simple; it satisfies all the “benchmark” requirements of stylized market facts including volatility clustering; is psychologically plausible; includes EMH as a particular case, so influences of different factors of deviations from EMH can be studied; robust as $N \to \infty$; takes into account transaction costs. There is also a mean-field mean-field version (a nonlinear RDS).

Points to notice: Not a book-order model; no binary interactions: interactions via sentiment, decisions triggered by price changes; no fundamental price; no different classes of traders; no internal dynamics: dynamics driven by the information stream.

Bad things: As this is not a binary interaction model, kinetic methods cannot be obviously applied. As this is not a book-order model, not clear how to apply MFG theory though this seems more promising. So difficult doing anything beyond “econophysics”: simulation and statistics analysis of simulation data. Lots of parameters that is not clear how to get empirically.

Introduction

I will suggest a construction for Bak–Sneppen type models which allows us to bring methods of denumerable Markov chain theory to bear on the question of finding bounds and approximating the threshold, as well as characterising the emerging fitness distribution in these models.

The classical model

The setup for the classical Bak–Sneppen model is a ring of N sites, that stand for ecological niches, each occupied by a species. Associated with each site is a fitness $x_k \in U[0, 1]$.

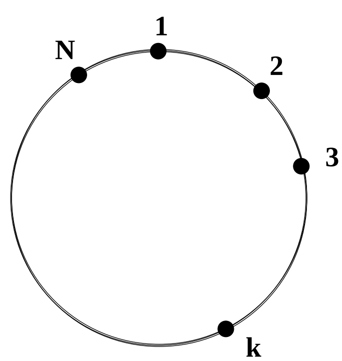


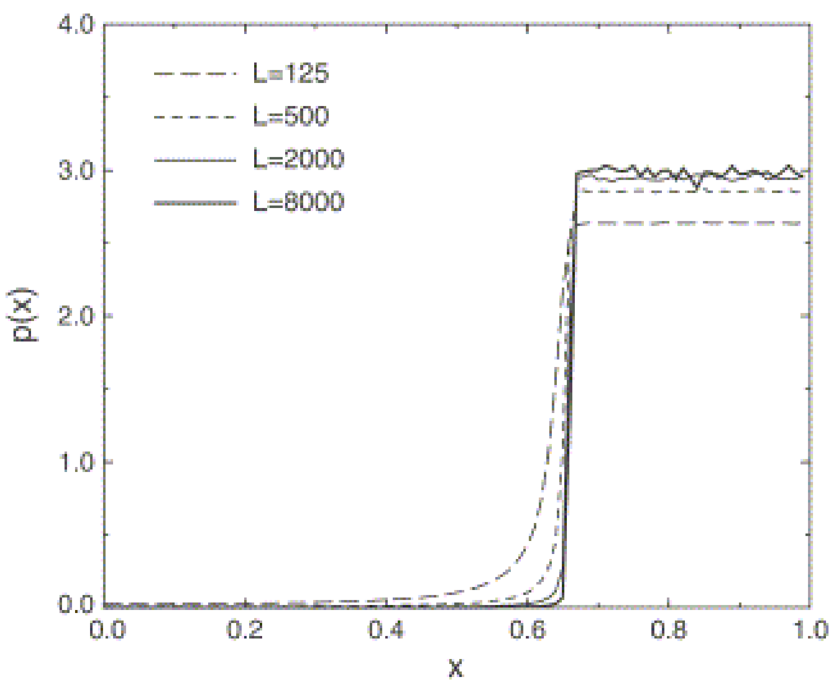
Figure 1: The Bak–Sneppen setup

The fitnesses evolve according to the following algorithm:

1. Find k_{min}: $x_{k_{min}} = \min\{x_k \,|\, k = 1, \ldots N\}$.
2. Replace $x_{k_{min}}$ and its two nearest neighbours by random numbers taken from $U[0,1]$.
3. Go back to 1.

Thus, we start with a distribution $\rho(0) = U[0,1]$, apply the algorithm and generate consecutively (samples of) distributions $\rho(1)$, $\rho(2)$, and so on.

The main question is: **what is the limiting distribution, $\rho(\infty)$?**



…**Numerically**, it looks as if there is a number which we will call s_{crit}, (in the classical Bak–Sneppen case $s_{crit} \approx 2/3$) and $\rho(\infty) = U[s_{crit}, 1]$. Analytically it is known that the mean of $\rho(\infty) < 1$.

The classical Bak–Sneppen model is analytically very complex, and a number of simpler models have been developed. Of interest is the anisotropic Bak–Sneppen model, which we will call aBS, but there are other models that are in a sense limits of the Bak–Sneppen model.

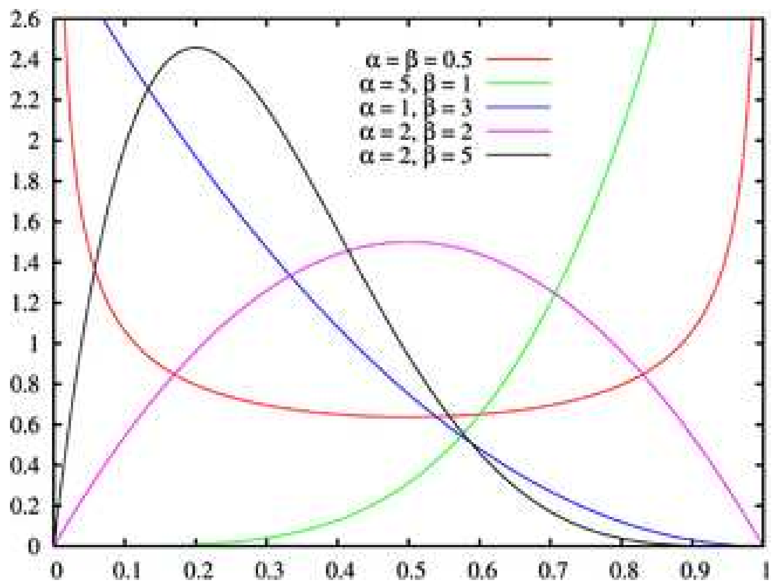
- The discrete Bak–Sneppen model: This model has been suggested by Barbay and Kenyon. Here $x_k \in \{0, 1\}$. The algorithm now is as follows:

    1. Pick a site with zero fitness at random, say x_j;
    2. Replace x_j and its two nearest neighbours by 1 with probability p and 0 with probability $1 - p$,
    3. Go back to 1.

- The Beta family: One question that has been neglected in the Bak–Sneppen literature is the question whether and how the fact that the initial conditions and new fitnesses are sampled from $U[0,1]$ influences $\rho(\infty)$. A nice way to investigate this question is to consider a wider class of probability distributions on $[0,1]$ that includes the uniform distribution, e.g. the Beta distribution, which has the PDF

$$\text{PDF} = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1},$$

Note (a) that this includes both bounded and unbounded and symmetric and asymmetric densities; the uniform distribution is $\alpha=\beta=1$, and (b) that taking $\alpha,\beta\to 0_+$ and keeping the mean constant allows us to recover the discrete Bak–Sneppen model for every p.



- n-Random Bak–Sneppen models: A useful simple model is the random Bak–Sneppen (rBS) model, in which fitnesses are and all the changes are in $U[0,1]$ and instead of changing the nearest neighbours of $x_{k_{min}}$, we change the fitnesses at n randomly chosen sites.
- And finally... The anisotropicBak–Sneppen model:

  In this model (aBS) , everything is as in the classical Bak–Sneppen model, except that at each step of the algorithm we change the fitnesses at $x_{k_{min}}$ **and at its right-hand nearest neighbour**.

Strategy and a construction

We use aBS a lot, as the computations for 2-rBS are trivial and for the classical Bak–Sneppen model quite laborious.

The strategy is as follows: assuming that there is a threshold, we will make a construction that will reduce the dynamics of the algorithm to a countable Markov chain.

Construction

Choose some $s \in (0, 1)$. Assume that initially one has a very large number N of fitnesses sampled from $U[s, 1]$. Now run the aBS algorithm, and consider the dynamics of M_k, the number of fitnesses in the interval $(0, s)$ after k iterations of the algorithm. We define

$$s_{crit} = \sup \left\{ s \,\Big|\, \frac{|\{k \leq K \text{ s. t. } M_k = N\}|}{K} \to 0 \text{ as } K \to \infty \right\}.$$

We would like to define a Markov chain, from which one can read off the limit $\lim_{k\to\infty} M_k$ when it exists.

The state space of this Markov chain needs to be discussed. Suppose we run the algorithm for k steps and have that $M_k = n$. This means that there is a string of n indices, $i_1, \ldots i_n$ such that $0 < x_{i_1} < \cdots < x_{i_n} < s$. We certainly have a Markov chain on I^N, but that is too large a state space. We want a Markov chain on **strings of indices**, ultimately of size N^2 in aBS and of size N^3 in BS.

Suppose we are doing aBS.

Then strings of length n belong in two classes: A_n if the fitness of right hand-side neighbour of the site with lowest fitness is in $[s, 1]$ or B_n if it is in $(0, s)$. Thus $[2, 5, 9, 3]$ is a B_4 string and $[1, 7, 12, 35, 36]$ is an A_5 string.

It is hard to define a Markov chain consistently on the whole set of strings of size less or equal to N ($O(N!)$); and a Markov chain cannot be really well-defined on equivalence classes: the probability of $[2, 4, 5]$ to lead to a B_2 string or a B_3 string is different from that of $[2, 4, 6]$, though both are A_3 strings. So the strategy is to average over classes of strings: the "empirical" Markov chain.

The construction can be easily implemented numerically: let C_1, C_2 be any two classes. Given s, L times generate real n strings, generate integer strings from them, and set

$$P(C_1, C_2) = \lim_{L\to\infty} \frac{|\text{a string of } C_1 \text{ is mapped by the algorithm into } C_2|}{L}$$

For BS, we need three classes of strings, say A_n, B_n and C_n, having, respectively, no, one and two nearest neighbours of the minimum fitness in $(0, s)$.

Tools from eager Markov chains

Once we have the Markov chain, it becomes clear that one can discuss the dynamics of M_k using the theory developed for probabilistic lossy communication channels: think about each fitness in $(0, s)$ as a message, fitnesses in $[s, 1]$ as a reservoir from which messages can come, and the algorithm as some disturbance (noise) mechanism that changes the state of the messages (N. Bertrand).

Let $\mathcal{M}$ be a (discrete time) Markov chain, i.e. a pair $(\mathcal{S}, \mathcal{P})$, where $\mathcal{S}$ is a (countable) state space and $\mathcal{P}$ is the transition probability matrix. Thus, if $r, t \in \mathcal{S}$, $\mathcal{P}(r, t)$ is the probability of moving from r to t in one move. If $T \subset \mathcal{S}$, $\mathcal{P}(r, T) = \sum_{t \in T} \mathcal{P}(r, t)$.

Let $\diamond T$ be the set of infinite paths in $\mathcal{S}$ that eventually visit T. Define $\Pr(r, \diamond T)$ to be the probability of reaching T from r.

Definition: $T \subset \mathcal{S}$ is an **attractor** for $\mathcal{M}$ if $\Pr(r, \diamond T) = 1$ for all $r \in \mathcal{S}$, Now let us restrict ourselves to a class of Markov chains that allow a particular decomposition: suppose that $\mathcal{S}$ can be decomposed into **levels** S_i:

$$\mathcal{S} = \cup_{i \in \mathbb{N}} S_i, \quad S_i \cap S_j = \emptyset \text{ if } i \neq j.$$

In other words, for each $r \in \mathcal{S}$, we can define

$$\text{level}(r) = \{j \,|\, r \in S_j\}.$$

[Note: in our Markov chain of strings formed by the fitnesses in $(0, s)$, the level of each string is simply the number of elements in that string.]

We set

$$\mathcal{E}(r) = \sum_{j=0}^{\infty} \mathcal{P}(r, S_j) \cdot j. \tag{1}$$

$\mathcal{E}(r)$ is the expected level for r to end up in under the dynamics of the Markov chain. We have

Theorem 1 *If $\mathcal{E}(r) \leq \text{level}(r) - \Delta$ for some $\Delta > 0$ and all $r \in \mathcal{S} \backslash S_0$, then S_0 is an attractor for $\mathcal{M}$.*

We illustrate how this theorem gives us a lower bound for threshold in aBS in BS.

Note that A_n (for $n \geq 4$, say) strings can create (among others) A_{n+1}, B_{n+1} strings, while B_n strings, under the action of the algorithm, can only give rise to A_n/B_n, A_{n-1}/B_{n-1} and A_{n-2}/B_{n-2} and certainly cannot go up a level. Similarly, in BS a C_n string cannot go up a level but can go to, say, A_{n-3}.

Thus, to find a crude estimate of s_{crit} we only need to consider A_n strings and compute the expected level after the algorithm is applied once to such a string.

Doing this, we have

Theorem 2 *In the case of aBS,* $s_{crit} \geq 1/2$ *(and for BS* $s_{crit} \geq 1/3$*).*

The proof is simple. Let us do the aBS case. Starting with A_n for any n (say, $n \geq 4$) we have that the probability of going up a level (either to B_{n+1} or to A_{n+1} is s^2, while the probability of going down a level is $(1-s)^2$. Hence to find a lower bound for the threshold, we only need to solve the equation

$$1 \cdot s^2 + (-1) \cdot (1-s)^2 = 0,$$

i.e. $s_{crit} \geq 1/2$.

[These numbers are **exact** threshold values for 1-random BS and 2-random BS, respectively, as you can prove by random walk methods.]

This is very far from known results: 0.66702 for BS and 0.7240 for aBS (Garcia and Dickman, 2005). With a bit of imagination we can do better. The equilibrium condition, say for BS, is

$$\sum_{n=0}^{N} P(n)r(A,n)[2s^3 + 3s^2(1-s) - (1-s)^3]$$
$$+ P(n)r(B,n)[s^3 - 3s(1-s)^2 - 2(1-s)^3]$$
$$+ P(n)r(C,n)[-3s^2(1-s) - 6s(1-s)^2 - 3(1-s)^3] = 0,$$

where $P(n)$ is the probability of being at level n, and the share of A_n, B_n and C_n strings in the set S_n of n-strings out of N has been denoted by $r(A,n)$, $r(B,n)$, and $r(C,n) = 1 - r(A,n) - r(B,n)$. These is true for every $s \leq s_{crit}$.

However, suppose we can characterise $r(A,n)$, etc. at criticality independently. If we could do that, we'd have an equation in s that would give us the value s_{crit}.

Let us consider the set of all n-strings for a fixed N, which we will call S_n. Set $\mathcal{A}_m^k = \frac{m!}{(m-k)!}$. There clearly are $\mathcal{A}_N^n$ n-strings in S_n. Also,

$$|A_n| = N\mathcal{A}_{N-3}^{n-1},\ |B_n| = 2N(n-1)\mathcal{A}_{N-3}^{n-2},\ |C_n| = N(n-1)(n-2)\mathcal{A}_{N-3}^{n-3}.$$

A reasonable criticality assumption is $P(n) = 1/N$,

$$r(A,n) = \frac{|A_n|}{|S_n|},$$

and similarly for the other fractions. So we are assuming uniformity over levels and in levels. This is wrong, but gives reasonable approximations.

It is not hard to prove that

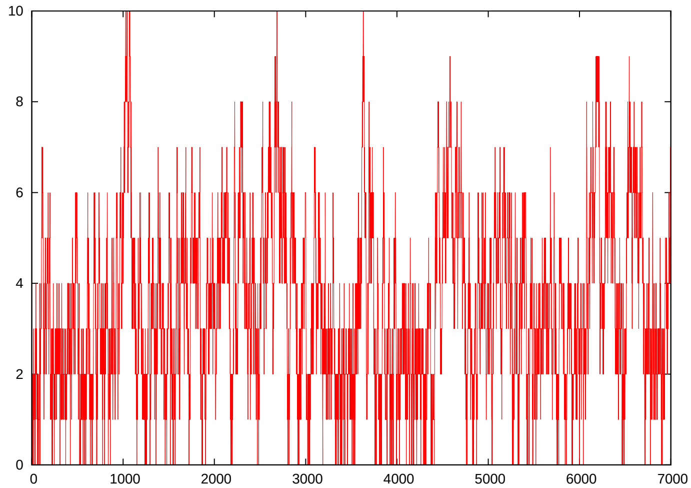
$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \frac{|A_n|}{|S_n|} = \frac{1}{3}.$$

The same is true for other sums, so what we have to solve then is

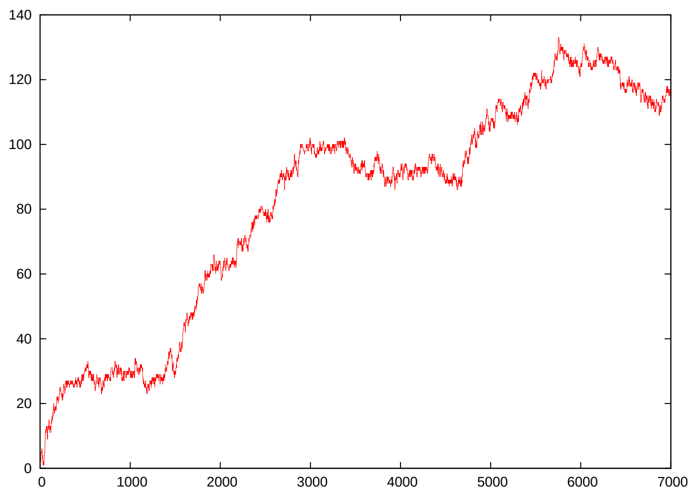
$$\begin{aligned}[2s^3 + 3s^2(1-s) - (1-s)^3] + [s^3 - 3s(1-s)^2 - 2(1-s)^3] \\ +[-3s^2(1-s) - 6s(1-s)^2 - 3(1-s)^3] = 0,\end{aligned} \tag{2}$$

which collapses to $3s - 2 = 0$. The equivalent result for aBS is $4s - 3 = 0$ (so unfortunately neither an upper not a lower bound, but pretty close!).

$s = 0.65$


10
8
6
4
2
0
0
1000
2000
3000
4000
5000
6000
7000

$s = 0.85$


140
120
100
80
60
40
20
0
0
1000
2000
3000
4000
5000
6000
7000

If instead of the uniform distribution, we use any other distribution and let $f(s) = P(x < s)$, then to find such an approximation of the threshold in the equivalent BS model all we have to do is to solve the equation $3f(s) - 2 = 0$. For example, amazingly for the Beta distribution with $\alpha = \beta = 1/2$ we also get $s_{crit} \approx 3/4$.

Is there a better a priori characterisation of criticality?

A coagulation-fragmentation framework

Let us assume that $s < s_{crit}$. Consider aBS for simplicity.

Suppose the Markov chain on $\{A_n, B_n\}$ has reached an equilibrium. Now let us denote by a_n, b_n the equilibrium probability that the system is in **one** of the states A_n or B_n, respectively so that

$$\sum_{n=0}^{N} (a_n + b_n) = 1.$$

Then (for $n > 2$; the cases $n = 0, 1, 2$ need special treatment), we have

$$a_n = \frac{s^2}{2} u_1 + s(1-s) u_2 + (1-s)^2 u_3$$

$$b_n = \frac{s^2}{2} v_1 + s(1-s) v_2 + (1-s)^2 v_3.$$

Here

$$
\begin{aligned}
u_1 &= \left[\frac{2(n-2)}{n-1} r(a, n-1) + \frac{1}{n-1} r(a, n)\right](a_{n-1} + b_n), \\
v_1 &= \left[\frac{1}{n-1} + \frac{2(n-2)}{n-1} r(b, n-1) + \frac{1}{(n-1)} r(b, n)\right](a_{n-1} + b_n) \\
u_2 &= \left[\frac{1}{n} + r(a, n) + \frac{n-1}{n} r(a, n-1)\right](a_n + b_{n+1}) \\
v_2 &= \left[r(b, n) + \frac{n-1}{n} r(b, n-1)\right](a_n + b_{n+1}) \\
u_3 &= r(a, n)(a_{n+1} + b_{n+2}) \\
v_3 &= r(b, n)(a_{n+1} + b_{n+2}).
\end{aligned}
$$

Here I denoted by $r(a, n)$ the (equilibrium) quantity $a_n/(a_n+b_n)$ and $r(b, n) = 1 - r(a, n)$. The logic of these equations is transparent, but they look useless if we want to take N large.

A possible work-around is as follows: We can create a system of difference equations by setting

$$\begin{aligned} a_n(k) &= \frac{s^2}{2}u_1(k-1) + s(1-s)u_2(k-1) + (1-s)^2u_3(k-1) \\ b_n(k) &= \frac{s^2}{2}v_1(k-1) + s(1-s)v_2(k-1) + (1-s)^2v_3(k-1), \end{aligned} \tag{3}$$

where now

$$u_1(k-1) = \left[\frac{2(n-2)}{n-1}r(a,n-1,k-1) + \frac{1}{n-1}r(a,n,k-1)\right] \times (a_{n-1}(k-1) + b_n(k-1)),$$

$$v_1(k-1) = \left[\frac{1}{n-1} + \frac{2(n-2)}{n-1}r(b,n-1,k-1) + \frac{1}{(n-1)}r(b,n,k-1)\right] \times (a_{n-1}(k-1) + b_n(k-1)),$$

$$u_2(k-1) = \left[\frac{1}{n} + r(a,n,k-1) + \frac{n-1}{n}r(a,n-1,k-1)\right] \times (a_n(k-1) + b_{n+1}(k-1)),$$

$$v_2(k-1) = \left[r(b,n,k-1) + \frac{n-1}{n} r(b,n-1,k-1) \right] \times$$
$$(a_n(k-1) + b_{n+1}(k-1)),$$

$$u_3(k-1) = r(a,n,k-1)(a_{n+1}(k-1) + b_{n+2}(k-1))$$
$$v_3(k-1) = r(b,n,k-1)(a_{n+1}(k-1) + b_{n+2}(k-1)).$$

If solutions converge to anything, we might hope it is the stationary distribution of our Markov chain. Note that we have to make a commitment what $r(a,n,k)$ is if $a_n(k)$ and $b_n(k)$ are both zero.

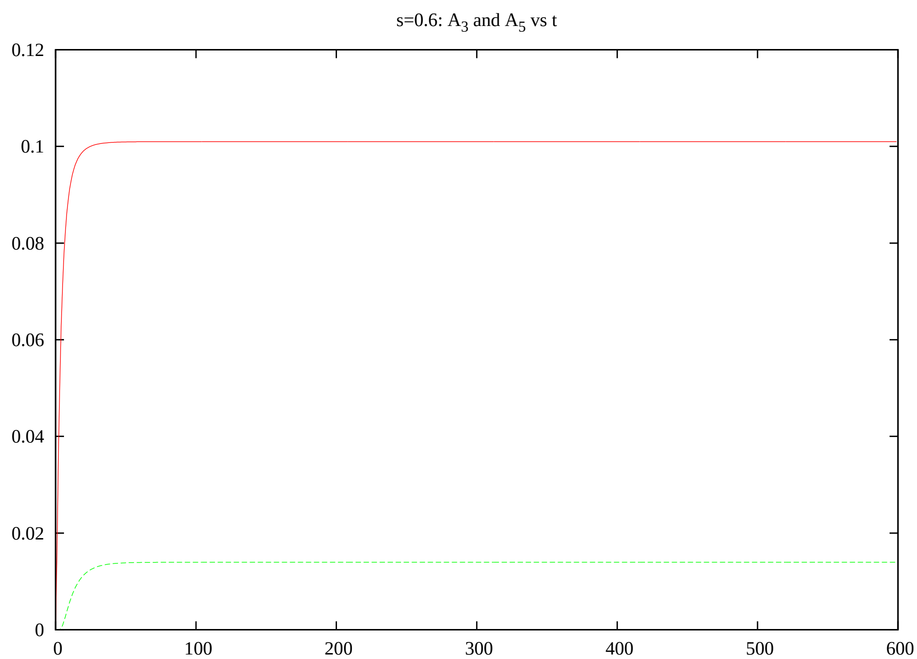
What does the DDS do

Note that the DDS is defined whether or not there is a stationary distribution $\{a_n, b_n\}$. While it is very difficult to *locate* the threshold s_{crit} numerically using it, it has two types of behaviour, which we show for $s = 0.6$ and $s = 0.8$.

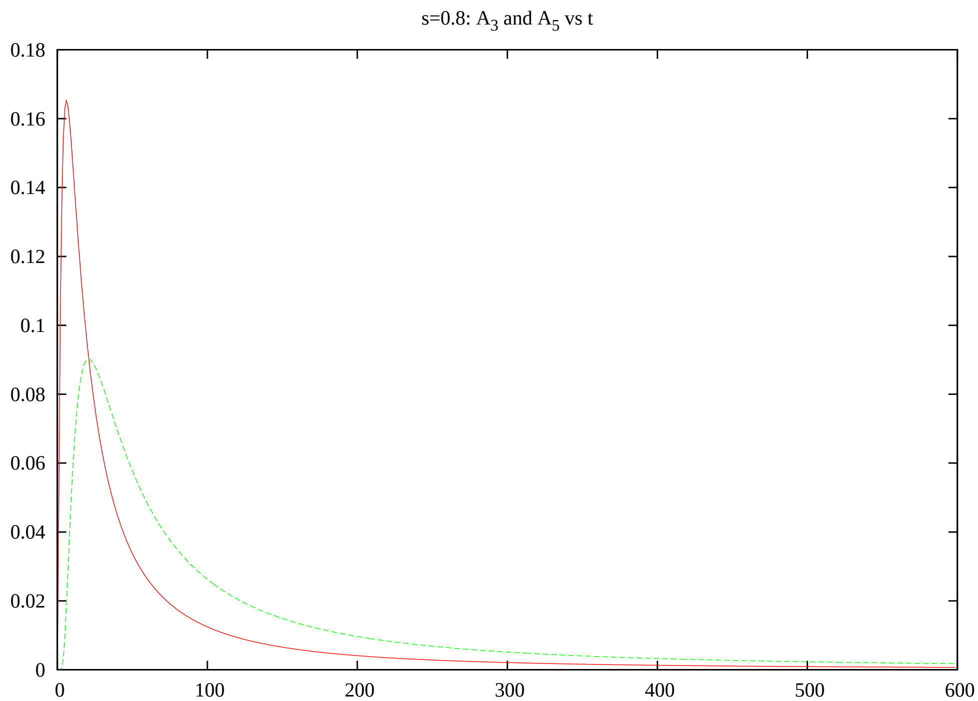
$s = 0.6$. This is a subcritical case, and one sees that all $a_n(k)$, $b_n(k)$ converge to a limit, as does the total number of fitnesses less than s,

$$M(k) = \sum_{n=0}^{\infty} n(a_n(k) + b_n(k)).$$

Here what $a_3(k)$ and $a_5(k)$ do:



$s = 0.8$**.** This is a typical supercritical case. More and more fitnesses hemorrhage into the region $s < 0.8$. Note the behaviour of $a_n(k)$!



Such behaviour is typical of many coagulation-fragmentation equations; in the supercritical case "mass runs off to infinity". It would be interesting to pursue this direction further.

Remarks

1. Note that we can predict the form of $\rho(\infty)$: if $\rho(0)$ has PDF $F_0(x)$ and $\rho(\infty)$ has PDF $F_\infty(x)$, the algorithm implies that

$$F_\infty(x) = \frac{1}{\int_s^1 F_0(u)\,du} F_0(x) H(x-s),$$

where $H(\cdot)$ is the Heaviside function.

2. It is true that the Bak-Sneppen model is not a model of anything in socio-economic sciences. But a similar model, which is, is very easy to formulate: as before, but now let x_i stand for risk aversion of a fund manager. I.e. she invests x_i of capital in bonds and $(1-x_i)$ is stocks. At each accountability period choose a return on stocks which must be a random variable, say s_k, (with positive drift) the return on bonds being constant, say r. So the i-th fund manager earns $s_k(1-x_i)+rx_i$. If she underperforms relative to the average earnings in the sector, sack and replace with probability p. This may explain the equity premium puzzle.

3. One can also make the rules in BS local: excise and replace the site i that minimises local fitness, i.e. $2x_i - x_{i+1} - x_{i-1}$, and its nearest neighbours.