Probability in Decline

Dean Brooks
Ekaros Analytical Inc.
Vancouver, B.C.
Brooks17@gmail.com
604-817-0734

What is nonlinear probability?

Application of the
maximum entropy principle
to predict gradual decline in
frequency of rare items with
increasing set size

Maximum entropy principle

- A probability distribution whose entropy is at least as great as that of all other members of a specified class of distributions
- Combines information constraint (minimizing what we expect to know about a system) with physical tendency of nature (systems wind up in this configuration).

Maximum entropy applied to nonlinear cases

1. If the set is incomplete and growing, neither average nor total entropy are necessarily constant
2. Increasing entropy for the set as a whole implies an increase in entropy for *later* elements.
3. Net entropy can be increased in this way if rare items become rarer. This behavior has been termed a ‘decline effect’.

An example from everyday life

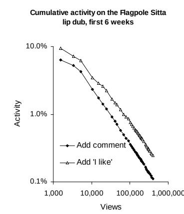
- Decline in participation in submitting comments on video-sharing websites like YouTube or Vimeo
- Initially, for every 10 views, one comment on average
- Average falls in very predictable fashion as the total set of views grows

Audience response to the ‘Flagpole Sitta’ lip dub video

Nearly perfect log-log relationship from 10,000 to 400,000 views

Participation in comments gradually falls from 2.3 % to 0.11 %

Behavior of visitors predictably different even from morning to afternoon of same day



Another practical application: H1N1 pandemic 2009

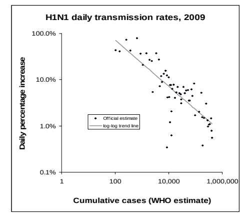
- H1N1 pandemic mortality and transmission rates based on cases in May were forecast to be high, turned out to be very low
- Nonlinear forecast model showed in June what the rates would fall to by October, but WHO continued to use linear assumptions

Official WHO data

The initial outbreak saw alarming transmission rates between 30 and 50 percent per day

By summer 2009, transmission fell to 5-10 % per day, often less

By the fall, transmission was well below 1 % and WHO had abandoned official case counts



The long search for a theory

- My initial hypothesis in the 1990’s was that decline required a certain minimum level of system complexity (it applied to people or biological organisms but not mechanical random-number devices).
- Experiments starting in 2003 showed otherwise: even very simple systems decline
- A lengthy literature search led to early ESP experiments and George Spencer-Brown

George Spencer-Brown (1923-)

- Wartime naval officer 1943-47
- Oxford mathematician and philosopher
- Colleague of Bertrand Russell, studied under Ludwig Wittgenstein
- Most famous for *Laws of Form* (1969)

ESP in the 1950’s

- Still semi-respectable at that time
- Papers would appear in *Nature*
- Blind, well-supervised experiments with good controls showed a baffling and persistent “decline effect”
- Correlations would start out high, then gradually fall off to below chance levels

Spencer-Brown’s hypothesis in *Probability and Scientific Inference* (1957)

- ESP is not reproducible, but the decline effect is.
- Assume that there is no ESP. Then these are large-scale experiments in probability and nothing more.
- Spencer-Brown’s own experiments using ‘chance machines’ without a human subject produce the same odd pattern.
- His hypothesis: perhaps these results are evidence of a subtle defect in probability theory itself.

‘Atomic’ and ‘molecular’ events

- Spencer-Brown referred to ‘atomic’ and ‘molecular’ events, atomic being ‘A’ or ‘B’ in this sequence, and molecular being ‘AB’ or ‘AA’ or ‘BB’ : AABABBAABAB
- The method laid out by Jaynes in 1957 calls for multiple layers of interpretation, each independent of the others (rate of occurrence of ‘AAA’ can decline or be different even though rate of occurrence of ‘A’ has not changed).
- Fundamental contrast with the approach of classical probability in which all ‘molecular’ events are strict permutations

Spencer-Brown’s experimental method

- Identification of ‘organic’ boundaries between data sets (e.g. change of ESP subject, new day, new machine)
- Comparison of first and last half (good for small data sets)
- Analysis by quadrants: First 25 percent of data set compared to last 25 percent

Classical expectation vs. actual result

- If experiments are well-controlled, there will be no bias toward fewer “rares” and trend graphs will tend to be flat
- The actual result: in hundreds of experiments, 80-90 percent show bias
- Bias is stronger as p(rare) << 1
- Bias persists longer as p(rare) << 1

Tools developed since Spencer-Brown

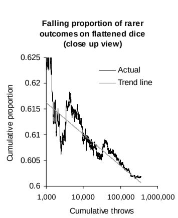
- Present the data as a cumulative log-log plot
- Show classical odds line if available
- Assess likelihood of deviation from classical odds line (most experiments have wound up in the range 0.1 > p > 0.01)

Experiment with flattened dice

I threw a set of four slightly flattened dice 100,000 times to see if rares would decline

The favored axis showed up 37.5 percent of the time, not 33.3

The favored axis increased its predominance over time



A large range of data sets

- Old ESP experiments
- Recent ESP experiments
- Spencer-Brown’s ‘chance machines’
- Modern random-number generators
- Casino games
- Dice, coins, toy roulette wheels

Weaknesses in standard testing

- Conventional tests work with relatively short blocks of data, often 1 kbit
- A surplus of long, repetitive strings (say 1111111111) cannot be detected using this approach. A shortage cannot be detected either.
- The decline effect encompasses the entire historical output of the machine, often hundreds of thousands of items. Testing would be expensive and time-consuming.

Entropy as the governing principle

- The most important clue is the consistent logarithmic decline
- Changes in the proportion of rares are steepest for low N
- Similar amounts of change for N, 10N, 100N events

Future prospects

This has enormous practical and theoretical potential in economics, biology, epidemiology, Web traffic analysis, sociology, military science.

Every field that uses probability will be affected.