Global errors in science: Traps of chance and prejudice

K. N. Ganeshaiah

Despite the awareness that the chances of a given ticket holder winning a lottery are dismally poor and sufficiently low enough to inspire any buying, several lakhs of people buy lottery tickets every month and lottery has been surviving as a successfull enterprise. Clearly the inspiration for the buyer does not come from those millions who have lost their rupees but from that occasional lucky winner who bags several lakhs of rupees. His smiling face printed in the magazines makes news and several more millions are induced to buying tickets; though they all are mostly bound to loose, again there is another lucky winner who helps maintain the buying chain and the lottery continues to survive. Thus unlike that of the proverbial chain, the strength of this lottery chain is reflected not in its weakest length (the million losers) but in the strongest link (the lucky winner).

Surprisingly, a similar process of occasional 'discovery' seems capable of sustaining a chain of false inventions in science on a global scale. Such a possibility has recently been pointed out by Bill Amos[1] from the University of Cambridge and discussed in the columns of *Nature*[2,3].

Assume that a certain interesting pattern (irrespective of how one wishes to define it[4]) has been proposed by a research worker in a symposium and that 20 of the participants are inspired to test it after their return. Even if the pattern does not exist, the laws of probability ensure that purely by chance one out of these 20 would find evidence for the pattern ($p < 0.05$). He would obviously be prompted to publish it because of his faith in the statistical significance of his experimental results and also because they are supporting what has already been held by others. The remaining 19 who could not get the 'expected' results may either feel less confident of their experimental protocols or more frequently do not find it interesting enough to publish the non-significant results. Occasionally some of them who do attempt to publish might find it hard to convince the referees against a pattern that is probably already making news.

Now assume that the paper published by the lucky scientist who got the results 'right' is read by a thousand research workers world over and that they would attempt to repeat the experiments. Again by the rules of probability, 50 of them would find the pattern to be true (< 0.05) and among them 10 would find it very highly 'significant' ($p < 0.01$). Even if only 30 of these 50 publish their results, it leads to a chain of the publications demonstrating a pattern that indeed does not exist!!

Chance and prejudice as traps

What happens next? Does this global chain of errors continue? Or would the self-correcting mechanism of science set it right? Certainly as a preliminary step, the scientific community is likely to celebrate this news and the scientists would begin working further on this pattern that has by now become 'established'. As I argued elsewhere[4], 'every pattern shown or demonstrated has the same effect as a miracle would have on the spread of a religion or the religious belief'. Such an error where a pattern is 'found' while it does not exist is termed Type I error in science and has the likelihood of not being noticed because of two reasons: one, the prejudice and the other the trap of chance.

First, scientists generally suffer from a prejudice of looking for only significant results – a syndrome that has been perpetuated by their incessant obsession to find patterns[4]. Consequently they have an instinctive desire to search for the existence of patterns, such that any work that does not find the pattern gets less or no attention both by the investigators and also by the rest of the scientific community. Therefore in situations where no pattern exists, 95% of the work demonstrating no pattern is likely to remain unpublished or if published might go unnoticed. This attitude of science is an important fitness component that has contributed for its survival[4].

Second, the scientific acceptance and approval of the pattern is innate in the statistical grammar or code the science has adopted for its own conduct: the results of the investigators who find the pattern significant has to be 'accepted' though it is likely that these might be only those statistically lucky investigators. Only a global observer with a bird's eye view of all the investigators could find that 95% of the workers have not been able to observe the pattern and have not reported this fact. But since the evaluation of the facts is not a well-coordinated global activity, rather is based mostly on individual cases, the 95% of the investigators are not going to add to the establishment (de-establishment) of 'the fact'.

There is one objection however: statistically if 50 scientists find the results significant by chance, 25 of them shall find the pattern in a positive direction and 25 in the negative direction such that the publication of both of these would end the matter and settle for a lack of pattern. But as Robert West[3] points out, scientists often seek trends in one direction (perhaps positive) and hence the chain process of reporting the results that 'conform' to the already held pattern would perpetuate.

Global errors as stochastic events

The operation of global errors of this nature is in fact a pure consequence of the stochastic events reinforced by a positive feedback process. Such process has been recently realized to be important in several fields[5]. It is consequent to the way in which the probability often drives a random system towards a pattern that appears deterministic.

Consider an industrialist planning to release two models of a new product to the market. Both the models have similar efficiency and hence he has no idea which of the two would make a breakthrough. Therefore he has formulated a simple strategy for further production of the model that customers like most: to begin with he releases just one unit each of the models and every time a unit of a model is sold in any of the outlets, he shall produce two units of that model and distribute. This way the production is directly linked to the rate of sale and the producer also guards against the overproduction of the model that does not sell well.

Accordingly, when the first customer buys one of the two models (say A), the industrialist releases two units of that model. Similarly the information on the model bought by the next customer is fed back to the producer who in turn releases two more units of that model to different outlets.

As this process goes on, and once the selling counters are flooded with the product, what would be the likely proportions of the two models in the market? Note that the marketing started off with one unit each of the models and that they cannot be discriminated by the customers as they do not have any specific differences in their customer appeal features; in fact their 'selection' of, or preference for, a model is a random process and hence would be in proportion to their availability in the market. Obviously we find it reasonable to expect that the frequencies of the models remain equal.

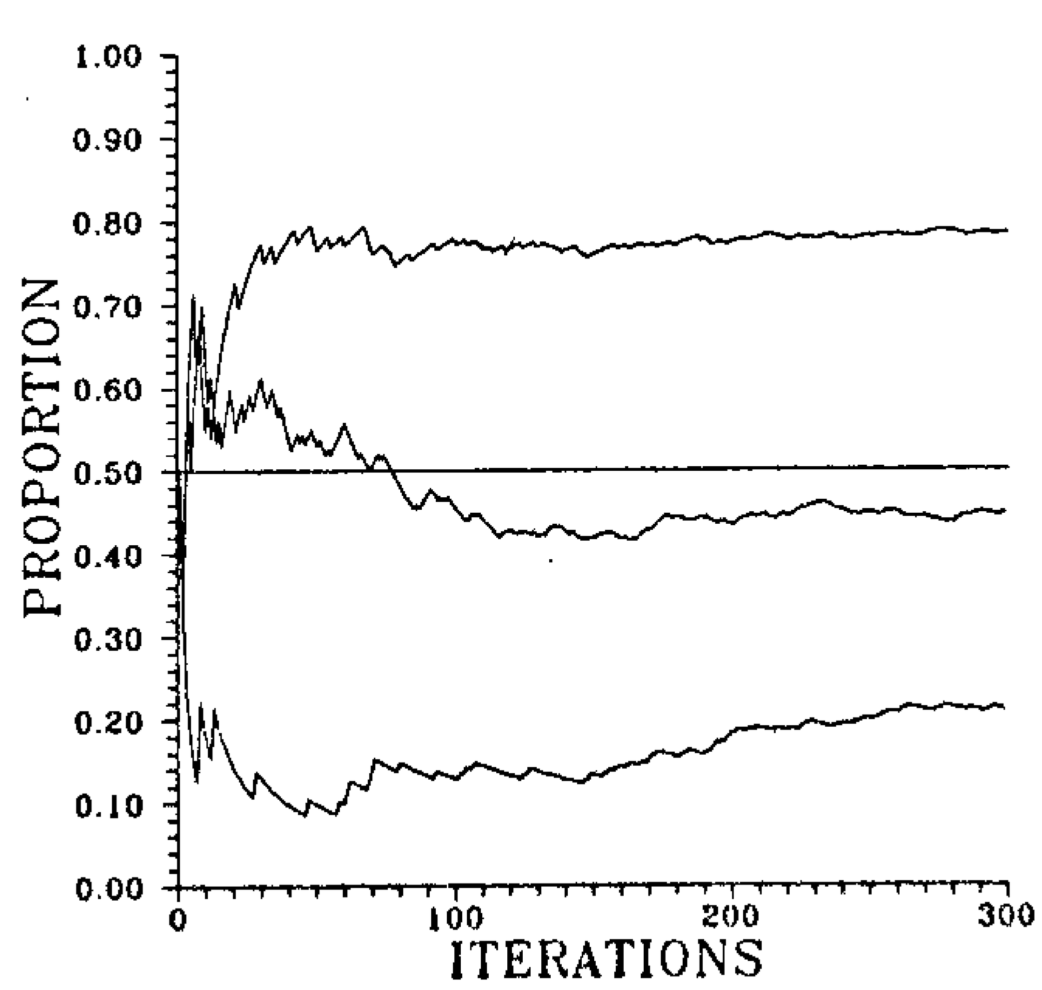

Figure 1. Frequency of red balls after every draw in three simulations of the Polya-Urn model. Note that the frequency begins with 0.5 as there are one each of red and white balls and hence it is generally expected that the frequency fluctuates around this (horizontal line). But as the process (draws; iterations) starts there will be an initial wiggle in the frequency which eventually stabilizes at some final proportion (p) in the range from 0 to 1. Summary of 2000 such simulations is presented in Figure 2.

But the mathematicians have shown that under such situations it is most likely that the market stabilizes with one of the models dominating over the other! Further, were the whole process to be repeated in another isolated island the chances are that the other model dominates the island market! In other words there is no way of predicting the fate of such models in advance.

The underlying stochastic events are well known to the mathematicians as the Polya-Urn process. The phenomenon can be visualized with a huge urn (market here) containing a red and a white ball (the models here); a ball is randomly drawn and returned to the urn with an additional ball of the same colour (as the industrialist does with the sold model) and the process is repeated. If one keeps track of the frequency of the two colours, initially, when there are just a couple of balls in the urn, the proportion of any colour shows heavy wiggles (Figure 1) because even a slight random bias in picking a particular colour can cause huge changes in the proportion. A chance favoured bias in drawing only red balls for instance, adds more red balls to the urn and increases the probability of drawing red balls further. However as the number of balls in the urn increases, this proportion converges to some limiting value p and as the urn gets filled with a large number of balls, the system becomes resistant to any random bias in picking the balls and hence gets stabilized. What is most interesting is that p can take any value between 0 and 1 (Figure 1).

I have simulated this process 2000 times and the results are plotted in Figure 2: clearly all values of p are equally

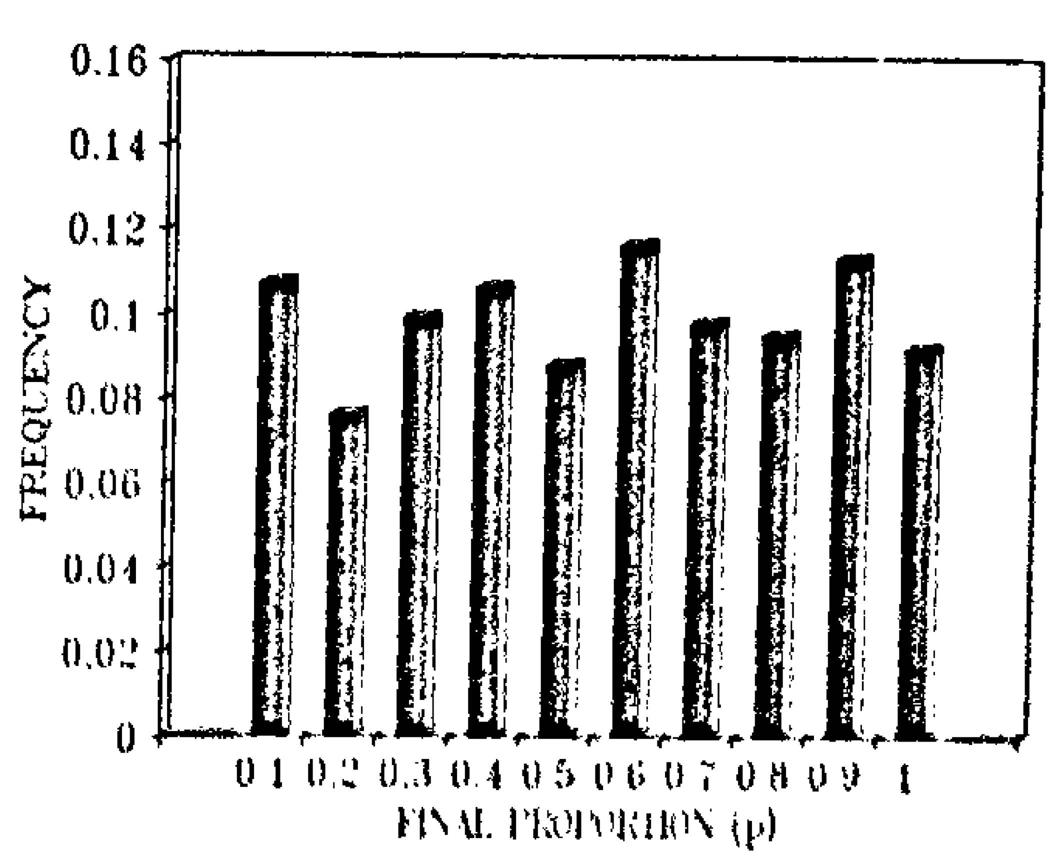

Figure 2. Frequency of the final proportions (p) of red balls in 2000 simulations. Note that all levels of p are equally probable and hence it is very difficult to predict the fate in advance.

probable! In other words, back to the marketing scenario, the extent to which a particular model dominates is purely a consequence of the history of the market and not a deterministic process. That science might frequently use these deterministic end results as the basis of investigation and attempts to probe for deterministic causes is another issue. A far more serious consequence of such stochastic process is the possible feedback chain of erroneous reports it might start off in science. As described earlier, such stochastic events based on the permissibility of the statistical codes of conduct adopted by the scientists might start a chain of false discoveries that might never appear to be the global errors in science. It appears it is very imperative to evaluate the extent to which our scientific information is clouded with such global errors generated purely by chance and nurtured by the prejudices of scientific community. **As Bulstrode[6] quips in *Nature*, 'The logical conclusion is, sir, your journal may be merely noise'.**

1. Amos, B., *Nature*, 1996, 379, 484.
2. Dunthom, D., *Nature*, 1996, **380**, 477.
3. West, R., *Nature*, 1996, **380**, 477.
4. Ganeshaiah, K. N., *Curr. Sci.*, 1995, **68**, 680–682.
5. May, R., *Nature*, 1976, **262**, 646.
6. Bulstrode, C., *Nature*, 1996, **379**, 765.

K. N. Ganeshaiah is in the Department of Genetics and Plant Breeding, University of Agricultural Sciences, GKVK, Bangalore 560 065, India and at Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India.