David's Musings: On password security, Republican candidates and predictability of economic crises

I recently refreshed my popular (general audience of non-mathematicians) talk on what mathematical probability says about the real world by introducing two new topics. I'm a big fan of back-of-an-envelope calculations, partly because I once wrote a book (...... *Poisson Clumping Heuristic*) consisting of one hundred such calculations, but more because I suspect that most of the real-world insights that mathematics provides can best be presented that way. If there were a Hall of Fame for back-of-an-envelope calculations then on prominent display should be the cartoon xkcd.com/936 which concludes *Through 20 years of effort, we've successfully trained everyone to use passwords that are hard for humans to remember, but easy for computers to guess*. The cartoon calculates that a password made of four common English words has more entropy than one made by substitutions of non-alphabet characters into an uncommon long word. As well as enabling an entertaining 5 minute introduction to the topic of Shannon entropy, I can then demonstrate the conclusion in real time, as follows. A popular talk demands audience participation, so before starting I asked the audience to supply 4 such words (and three numbers between 10 and 50, for the next topic) and wrote them on the board. At this point, I go to an online password strength checker, type in the four words, for instance, clockparrothappylevel, and invariably this is deemed a "strong" password. Next, I pull out the paper on which my hosts gave me some impossible-to-remember password to access their wireless network, type it in and invariably this is deemed a "weak" password. The audience looks impressed!

My second new talk topic concerns the almost finished two-year US race to determine the 2012 Republican Presidential Nominee. Almost all commentaries on the race remark that there have been an unusually large number of candidates (Sarah Palin, Rick Perry, Newt Gingrich, Michele Bachmann, etc.) whose popularity has risen and then dramatically fallen. But is it really *unusually* large? There are two ways one might think about this question. Opinion polls ask who a voter supports right now. There is no mathematical theory concerning how rapidly people can change their opinions, so studying the question via opinion poll fluctuations over time would involve purely empirical comparisons with data from previous electoral cycles. But what mathematics does say is that the probability of a specified future event happening, give the information known at time t, must evolve as a martingale. Though in general one cannot observe probabilities, in this context we can look at the Intrade prediction market (http://www.intrade.com/) where one can buy and sell contracts on candidates. A market price of 40 reflects a consensus probability of 40% that the candidate will be the nominee; and we can observe how the prices have fluctuated over time. Such markets are interesting because theory -- the "efficient market hypothesis" that market prices do indicate true probabilities -- gives testable predictions. In the context under discussion, an interesting mathematical prediction is:

for any price x, if each candidate's initial price is below x, then the expected number of candidates whose price ever exceeds x equals $100/x$.

At this point in the talk, I show the data on maximum prices for each candidate.

Romney 98; 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.

I can then "test" the prediction using the numbers x the audience gave me before the talk. The prediction for "over 22" is 100/22, and so on, and you can see the data matches the predictions pretty well. One could have a lengthy discussion of what this signifies -- for instance, that the smart money is not unduly influenced by fluctuating opinion polls. To me, the bottom line is that the only statistically unusual feature of the campaign in this sense has been that it started without any very prominent candidate.

A third talk topic is one that refreshes itself. Each year since 2006 the OECD has produced a "global risks report" for the World Economic Forum annual meeting in Davos, containing a graphic showing perceived likelihood and economic impact of 36 potential "risks", in categories such as economic, geopolitical, environmental, societal, technological, with the list changing somewhat from year to year. The latest report is available online (http://www.weforum.org/reports), and I show and discuss it briefly. But my main aim in the talk is to investigate how accurate were the old assessments. In particular, how predictable was the global late-2000s financial crisis (http://en.wikipedia.org/wiki/Late-2000s_financial_crisis), as Wikipedia calls it? I show the report written in mid-2007, at which time there were concerns about the worldwide boom in house prices, and some concerns about US subprime mortgages, but nothing dramatic had happened in other markets. The five most serious risks, combining likelihood and impact, were perceived at that time to be

Asset price collapse -- Oil price shock -- China economic hard landing -- Inter-state and civil wars -- Breakdown of civil informational infrastructure.

Given that these 5 risks were assessed to have 10--20% likelihood and that the first one actually occurred (albeit with substantially more than predicted severity), this OECD assessment is surely as good as one could hope for. Note also that the "oil price shock", assessed as second most serious, seems in retrospect to have been about to occur in 2008 but was overtaken by the asset price collapse.

My point -- surely self-evident to Bernoulli Society readers, but not to the outside world -- is that instead of deterministic "forecasts" of uncertain aspects of the future, one should make explicitly probabilistic assessments of different possibilities. It is easy to casually assert that quantitative predictions of the future, from Malthus on, have mostly turned out to be much less accurate than their authors supposed, but these last two talk topics give some encouragement that probabilistic assessments may be turn out more reliable.

David Aldous, Berkeley

Editor's note: This is the fifth installment of a regular opinion column.

Past Conferences, Meetings and Workshops

Snapshots of Stochastics Frontiers at 180 Degrees (SF-180)

This international symposium was held in Helsinki on 8-9 December 2011, to celebrate the 60th birthdays of three Finnish probabilists Esa Nummelin, Paavo Salminen and Esko Valkeila. The invited speakers, all long-term collaborators of the triplet, were:

- Christian Bender (Saarland U)
- Andrei Borodin (Steklov, St. Petersburg)
- Kacha Dzhaparidze (CWI, Amsterdam)
- Alexander Gushchin (Steklov, Moscow)
- Ingemar Kaj (Uppsala U)
- Takis Konstantopoulos (Uppsala U)
- Andreas Kyprianou (U Bath)
- Yuliya Mishura (Taras Shevchenko National U, Kiev)
- Pierre Vallois (Henri Poincaré U, Nancy)
- Lioudmila Vostrikova (U Angers)
- Marc Yor (Paris VI U)

In addition, Esa, Paavo and Esko gave special talks recalling their favorite theorems along their career in a special honorary session chaired by Elja Arjas. The opening speech of the meeting was given by Mats Gyllenberg. The evening program of the symposium included a traditional Finnish Stochastic Sauna evening, and was followed by a dinner spiced up with an oriental dance show by Saara Lehto and a juggling performance by Harri Varpanen.

From left to right: Esa Nummelin, Paavo Salminen, and Esko Valkeila, three Finnish probabilists celebrating their 60th birthdays at SF-180

The meeting was organized by Lasse Leskelä (chair), Dario Gasbarra, Göran Högnäs, Ari-Pekka Perkkiö, and Tommi Sottinen. The meeting was sponsored by the Finnish Doctoral Program in Stochastics and Statistics (FDPSS), Aalto University, the University of Helsinki and the Åbo Akademi University.

For more information, please see: http://web.abo.fi/fak/mnf/mate/gradschool/homepage_files/SF180.html

Lasse Leskelä (Jyväskylä)

XII Latin American Congress of Probability and Mathematical Statistics

The 12th edition of the Latin American Congress of Probability and Mathematical Statistics (CLAPEM), endorsed and co-sponsored by the Latin American chapter of the Bernoully Society, SLAPEM, was organized by Universidad de Valparaiso, Pontificia Universidad Católica de Valparaiso, Universidad de Santiago, Pontificia Universidad Católica de Chile and Centro de Modelamiento Matematico, Univesidad de Chile, in Viña del Mar, Chile. The meeting was held last March, 26th--30th, in the Hotel O'Higgings. Around 230 participants from several continents contributed to this successful, diverse and multinational meeting, confirming that the CLAPEM has become the main periodic scientific meeting on statistics and probability in Latin America.

Activities included 14 plenary and sub-plenary talks by Madalin Guta (The University of Nottingham, UK), Michael Jordan (University of California, Berkeley, USA), Steven Lalley (University of Chicago, USA), Yanyuan Ma (Texas A & M University, USA), Fabio Martinelli (University of Rome, Italy), Carl Mueller (University of Rochester, USA), Victor Pérez-Abreu (Centro de Investigaciones Matemáticas, Mexico), Marina Vannucci (Rice University, USA), S.R.