You shall live adventurous times

Momentous times for statisticians at times of need and tumult

Commencement Speech at Berkeley, Statistics, May 19, 2014

Andrea Saltelli
European Commission, Joint Research Centre,

You shall live adventurous times;

NEED: Statistics is badly needed at the hearth of the scientific enterprise

Poor statistical practice – or total neglect of statistical practice altogether, is very common. Two preceding commencement speeches (John P.A. Ioannidis and Nassim N. Taleb) delved into that. Allow me to continue this tradition briefly by using use one illustration from my craft: the way sensitivity analysis is run on mathematical models used to support important policy decisions. [This is the only slightly technical part of the talk]. In most mathematical models the robustness of the inference is tested by moving one factor/assumption at a time. This seems pretty innocuous … isn't it?

The point is: nobody would ever dream to do something so silly in natural experiment design; not since Charles Peirce developed methods for experimental design in the XIX century.

This is but one of the many instances of the curse of dimensionality. Think of a sphere inside a cube touching its faces; then imagine moving from the usual three dimensions of our ordinary space to many more… if you can! (Russian mathematician Kolmogorov apparently said he could find its way in a four dimensional cube…).

Though the cube is already larger than the sphere in three dimensions, when increasing the number of dimensions the volume of the sphere becomes tiny compared to the volume of the cube. Moving one factor at a time from within the sphere you always get as far as the surface of the sphere; it does not allow you to explore the cube's corners. Multidimensional cubes resembles hedgehogs … they are full of corners! Bottom-line: moving one factor at a time only scratches the real problem … not to mention that it does not allow you to see the interaction effects (Saltelli, A., Annoni, P., How to avoid a perfunctory sensitivity analysis, Environmental Modeling and Software (2010), 25, 1508-1517; Sam Savage, 2009, The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty, Wiley).

TUMULT FROM NOW ON: Times ripe with controversy. Not just climate and GMO but practically every issue involving science/[=statistical facts], from bees and pesticides [bitter dispute over whether the developed world's most popular pesticides are causing an ecological

catastrophe] to shale gas fracking, from the cooling liquid of Mercedes-Benz [for the German the ozone friendly liquid mandated by the EU guidelines was flammable, the French retorted that it was only flammable at high temperature …] to endocrine disruptors, from the fate of children raised by gay parents to the true long term cost of citizenship for illegal migrants now in America, …). ; Most controversies about facts relevant to society and its environment have statistics at their hearth.

Times fraught with normative hurdles, where more and more issues become 'wicked', meaning by this deeply entangled in a web of hardly separable facts, interests and values … mention Galileo (first) and Boyle (later) strenuous effort to create 'matters of fact'… (Steven Shapin & Simon Schaffer , 1985, Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life, Princeton, 2011 Edition)

Times where the language is often unceremonious. Mention the media/blogosphere norms '*Beware the rise of the government scientists turned lobbyists […] From badgers to bees, government science advisers are routinely misleading us to support the politicians' agendas*'. (George Mombiot, The Guardian, Monday 29 April 2013). You have heard similar language in exchanges in the US among opposing factions, e.g. on EPA regulations or between opposing factions on climate, GMO etc … more later

Times where uncertainty is routinely fabricated or concealed by opposing factions.

Fabricated. Mention book *Merchants of Doubt* of Naomi Oreskes and Erik Conway (Oreskes, N. and Conway, E.M. (2010) Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming, Bloomsbury Press, New York.)

Concealed. Mention Uncertainty and Quality in Science for Policy by Silvio Funtowicz and Jerry Ravetz: the concept of 'GIGO' or 'pseudo-science', "where uncertainties in inputs must be suppressed lest outputs become indeterminate". The same concept exists in econometrics (Leamer, E., 1990, 'Let's take the con out of econometrics', American Economics Review, March 1983, 73, 31–43; Kennedy, P. (2007) A Guide to Econometrics, 5th ed., p.396, Blackwell Publishing, Oxford).

Times where the call for quality can be seen to cut both ways.

> Quote from Secret [sic] Science Reform Act of 2014 – (already telling EPA that their science is secret doesn't sound as a compliment):
>
> '[…] to prohibit the Administrator of the Environmental Protection Agency (EPA) from proposing, finalizing, or disseminating a covered action unless all scientific and technical

information relied on to support such action is specifically identified *and publicly available in a manner sufficient for independent analysis and substantial reproduction of research results*.' Here the normative me and the technical me are split like Jim Carrey in Me, Myself & Irene …

Times were statistics is at the core of the storm

Statistics' rule. According to historian of science Ian Hacking (Hacking, I., 1990, The Taming of Chance, Cambridge: Cambridge University Press) 'Probability' won an epistemological war [epistemology studies the way we go about knowing, how do we decide that we know what we know] in the nineteen century. 'Probability' became king in adjudicating the credibility of evidence. We look at facts mostly through the lenses of statistics (in the previous centuries – up to the enlightenment - chance was equated with superstition). According to Hacking the victory of probability was metaphysical, epistemological, logical and ethical, leading to the 'imperialism of probability'…

The book of Hacking is a 'must read' & 'page turner' for any statistician; he tells how the world became 'numerical' between the XVIII and XIX centuries, before the 'big data' of our generation; the fascinating story of Leibnitz '*philosophical godfather of Prussian official statistics*' to the Prince Frederik of Prussia 1700; 56 categories to 'measure the power of a state' (the first scoreboard; e.g. number of marriageable girls, able bodied capable to carry arms, diseases, child mortality…); Leibnitz's first proposal for a statistical office … already in 1745 Jews are being treated as a separate category and counted in Prussian statistics … to finish a discussion – with reference to Laplace, Peirce and many others, - on the double face of probability, as both 'frequency' and 'degree of belief', and the note that at young age one is frequentist (nominalist he says) and later in life one becomes 'realist' (we would say Bayesian).

In the hearth of the storm. Title 'How science goes wrong' splashed on the cover page of The Economist (Oct 19th 2013): 'Science *still commands enormous—if sometimes bemused—respect. But its privileged status is founded on the capacity to be right most of the time and to correct its mistakes when it gets things wrong. […] The false trails laid down by shoddy research are an unforgivable barrier to understanding*.'

Maybe the most damaging example in this respect is the spreadsheet excel error that costed economists Carmen Reinhart and Kenneth Rogoff their reputation. The story was that trespassing a 90% threshold of public debt (as pct of GDP) countries would stop growing … no such threshold existed but Reinhart and Rogoff's work was used by 'austerians'. The error was spotted by a graduate student at the University of Massachusetts

Reproducibility debacle. '*… the majority of preclinical cancer papers in toptier journals could not be reproduced*.' (Begley, C.G., 2013, Reproducibility: Six red flags for suspect work, Nature, 497, 433–434)

Quote from the article (Ioannidis): *Why Most Published Research Findings Are False*. On one hand a statistical problem (the issue of false positive and false negative as the real issue for science's purported lack of reproducibility), on the other hand '*a research finding is less likely to be true when* […] *there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance*.

According to historian of science Philip Mirowski the normative dimension is predominating. His narrative: after the eighties neoliberal ideologies succeeded in decreasing state intervention in the funding of science, which became increasingly privatized. Knowledge as a monetized commodity replaced knowledge as public good. In house science labs of major corporations were closed and research outsourced to universities which became more and more looking as commercial outfits. Then research ended up outsourced again to even cheaper contract-based private organizations working on a short leash. (Mirowski, P., 2011, Science-Mart: Privatizing American Science, Harvard University Press).

Mention that the point of science's legitimacy challenge was made by philosophers, e.g. Jean-François Lyotard in the seventies.

Times where sides have to be taken (Like in Battlestar Galactica ;-)), hypocrisy resisted and humility exerted (Designs on Nature: Science and Democracy in Europe and the United States, Sheila Jasanoff, 2007, Princeton University Press).

- **Hypocrisy:** Let's avoid the illusion of speaking truth to power

  Statistics for policy: three models (Boulanger, P-M., Political uses of social indicators: overview and application to sustainable development indicators. International Journal of Sustainable Development, 10 (1,2):14-32, 2007).

    A rational-positivist model for the use of indicators and policy (good quality statistics underpin good policies)

    Discursive-interpretive model (statistics contribute to a process of framing of and focusing on an issue among the many competing for public's attention)

Strategic model (statistics is used by parties competing for a given constituency – a famous example the debate between Al Gore and George W. Bush – when apparently Bush's quip: '*I'm beginning to think not only did he invent the Internet, but he invented the calculator. It's fuzzy math*' won him the day and perhaps the presidency).

Again, is it possible to disentangle evidence based policy from policy based evidence? Guess, which were the first R&D statistics ever to be collected? Well before patents and citations and university rankings were ever looked at, someone looked at the number of offspring of reputed living scientists. Why? To prove eugenics. The idea was that intelligent people reproduce less so mankind is condemned to stupidity. (Benoît GODIN on eugenics and the birth of R&D stats: The Culture of Numbers: From Science to Innovation, INRS, Montreal, Canada, Communication presented to the Government-University-Industry Research Roundtable (GUIRR) US National Academy of Sciences, Washington, May 21, 2010). Statistics has its share of responsibility (perhaps a large one) in Eugenics' fortunes, see works of early statistician Francis Galton (Hacking's book quoted above).

- **Humility:** Let's resist the deficit model (if only politicians, the citizenry, … understood better science/statistics then progress would be achieved)
- Let's ask the right question – fighting type III errors. The example of GMO.

'*Montpelier is America's only McDonald's-free state capital. A fitting place, then, for a law designed to satisfy the unfounded fears of foodies*'; In another piece always on the Economist: *'Just ask about genetically modified crops, declared safe by the scientific establishment, but reviled as Frankenfoods by the Subarus-and-sandals set.'*(The Economist, May 10th 2014, Genetically modified food. The little state that could kneecap the biotech industry).

The PABE study and the real concern behind GMO (Final Report of the PABE research project funded by the Commission of European Communities, Contract number: FAIR CT98-3844 (DG12 - SSMI), December 2001).

Why do we need GMOs? What are the benefits?

Who will benefit from their use?

Who decided that they should be developed and how?

Why were we not better informed about their use in our food, before their arrival on the market?

Why are we not given an effective choice about whether or not to buy and consume these products?

Do regulatory authorities have sufficient powers and resources to effectively counter-balance large companies who wish to develop these products?

- **Humility.** Let's resist scientific/statistical hubris; risk is different from uncertainty (which is different from indeterminacy); you may compute something with the craft of Bayesian calculus but be careful to present this as solution to the problem of genuine uncertainty.

  Quote from Mary Douglas: '*Every choice we make is beset with uncertainty. This is the basic condition of human knowledge. A great deal of risk analysis is concerned with trying to turn uncertainty into probabilities. What seems to be in each case a purely technical exercise quickly becomes one that rests directly upon the philosophical foundations of inference*.'

  '*The effort of turning uncertainty into probability, i.e. into "known", calculable risk is pursued even when dealing with complex systems, made of tightly coupled components. Charles Perrow coined the original expression "normal accidents" to signify that failures are inevitable in such systems including those carefully designed (such as nuclear power plants), which one would expect to be well understood and controllable.*' (De Marchi, B., From the Taming of Chance to the Rhetoric of Uncertainty', POLITEIA, XXVI, 97, 2010. ISSN 1128-2401 pp. 3-10).

  All of the above did not convince you? This debate - quite some time ago - opposed Frank Knight and Maynard Keynes on one side and Frank Ramsey, and Jimmie Savage (the father of Sam Savage) on the other.

  John Kay, a British economist: "*For Keynes, probability was about believability, not frequency. He denied that our thinking could be described by a probability distribution over all possible future events, […] . In the 1920s he became engaged in an intellectual battle on this issue, in which the leading protagonists on one side were Keynes and the Chicago economist Frank Knight, opposed by a Cambridge philosopher, Frank Ramsey, and later by Jimmie Savage, another Chicagoan. Keynes and Knight lost that debate, and Ramsey and Savage won, and the probabilistic approach has maintained academic primacy ever since. A principal reason was Ramsey's demonstration that anyone who did not follow his*

*precepts – anyone who did not act on the basis of a subjective assessment of probabilities of future events – would be "*Dutch booked*"."* (Kay, J., August 14, 2012, Financial Times, The other multiplier effect, or Keynes's view of probability).

- At times when we seem to have methods to formally compute risks, the answer remains elusive. *Chernobyl - how many died?* Fifty persons died on the spot fighting the disaster, but if the 'Linear No Threshold' model holds, death could be in the range of one hundred thousand, or even more, and of course the debate in non-neutral, … Greenpeace versus the United Nations Scientific Committee on the Effects of Atomic Radiation. (Jim Green - Nuclear Monitor: Chernobyl - how many died? 2014, 26th April 2014).

- Policy based evidence is the flip side of evidence based policy (Strassheim, H., and Kettunen, P., Evidence & Policy, **10** (2), 259-77, 2014), and it is impossible to extricate the two, exactly as it is impossible to extricates facts from value when operating at the interface between science (statistics) and policy.

What then?

One must be brave to suggest a recipe to tackle all of the above. I am encouraged by the editorial of James Zidek, a reputed statistician at the University of British Columbia (Zidek, J., 2006, Editorial: (Post-normal) statistical science, Journal Royal Statistical Society A, **169** (1), 1–4) suggesting post normal science to statisticians.

He notes that in the Royal Statistical Society's code of conduct *'little specific guidance concerns conduct in adversarial proceedings'.* He looks instead at post normal science (PNS), a concept developed by Silvio Funtowicz and Jerry Ravetz in the early nineties (Funtowicz, S, Ravetz, J, 1993, Science for a post-normal age, Futures 25, 7, 739–55), to tackle *'[…] the intersection of values, public policy and science […] taking place in public view, often adversarial, having a multiplicity of stakeholders and involving concern for accountability (the bottom line!).*

Zidek notes that '*Interest groups have well-honed strategies for exploiting PNS's uncertainty: raising doubts (e.g. confounding and measurement error); asking the wrong questions; making numerous freedom-of-information access requests and using other forms of harassment; re-*

analysing the data, perhaps with different models or assumptions. There is plenty of scope here for a postnormal statistical scientist!'

Plenty of scope then to learn to work in post-normal science (PNS) settings and habits, i.e. with/on scientific inquiry that impacts on policy and society; here the facts are uncertain, values in dispute, stakes high and decisions urgent (Funtowicz & Ravetz, 1993) . The Chief Science Advisor of New Zealand Peter Gluckman just adopted this stance (Gluckman, P., 2014, The art of science advice to government, Nature, **507** , 1 6 3-165).

> *I have come to understand that the primary functions and greatest challenges for a science adviser are providing advice not on straightforward scientific matters, but instead on issues that have the hallmarks of what has been called post-normal science.*
>
> *These issues are urgent and of high public and political concern; the people involved hold strong positions based on their values, and the science is complex, incomplete and uncertain. Diverse meanings and understandings of risks and trade-offs dominate.*

What to measure, how to model, foremost the framing of the analysis, all is the result of a societal negotiation (*social inquiries* according to John Dewey, rather than knowledge discovery as in traditional science). In PNS Objectivity gives way to Inter-subjectivity, peer review gives way to extended peer review (stakeholders involvement – the concept of socially robust science), mere quality control gives way to accountability, the process generating the facts and their interpretation becomes crucial to the negotiation (whose facts, whose utility, who is the storyteller...).

We may find ourselves engaged in activities such as arbitrating conflicting statistical facts or working at places like Ioannidis' METRICS (Meta-Research Innovation Centre launched at Stanford), policing scientific production …

Avoid stealth advocacy. Exert ourselves in 'science brokering' (Roger A. Pielke Jr., 2007, The Honest Broker: Making Sense of Science in Policy and Politics, Cambridge University Press).

Back to Zidek's JRSS Editorial: *A partial solution lies in ensuring that statistical education is sufficiently broad to acquaint statistics graduates with the challenges that are presented by PNS while equipping them with the enhanced skills that are needed to cope with them and even selectively to take advantage of the opportunities.*

In particular, the statistical consulting sequence that is commonly found in statistics graduate programmes might be expanded to include multidisciplinary meetings where a multiplicity of legitimate views are presented in an adversarial context.

Well, in the end, to do our job properly – it seems – we have to go back and study some rhetoric; not a bad conclusion after all.