RMM Vol. 3, 2012, 19–21
Special Topic: Statistical Science and Philosophy of Science
Edited by Deborah G. Mayo, Aris Spanos and Kent W. Staley
http://www.rmm-journal.de/

Stephen Senn

Names and Games

A Reply to Deborah G. Mayo

I started paying attention to Bayesian claims some years ago when I found that students of mine were being criticized for not using Bayesian methods. If only they had used a Bayesian approach to what they were doing it would have been so much better etc. When I investigated further I discovered that many of these critics were not very clear about what a Bayesian was and I have not infrequently found myself in the position of having to explain to those who fly under the Bayesian flag of convenience what is they have signed up to. Just to give one example, one commonly encounters the misunderstanding that Bayesian analyses can get by with uninformative prior distributions. This is, of course, not true. A lot of what goes by the name of Bayesian analysis formally uses vague priors but as Jimmy Savage once pointed out a model should be as big as an elephant and in that he was being too modest: Any Bayesian model has to be as big as the universe itself and anything that is not formally in it has a highly informative prior attached to it that its effect is negligible.

This may seem like a linguistic quibble, but it's not. I once argued with an eminent Bayesian (but got nowhere) that it was impossible to insist that all meta-analyses be random effect meta-analyses unless one was prepared to be serious about several aspects of formulating the problem. What was one trying to predict: the effect in future patients or in future trials? How did one apply the method when one only had one trial? In what did genuine replication in the space of all possible trials past present and future consist? What did one have to believe about the joint prior distribution of effects and variances?

However, my day job has been at various times in my career and is again, statistical analysis for real. In my days working in the pharmaceutical industry I used to warn the statisticians in my group that they had to be able to answer 'Yes' to the following question: 'Am I seriously interested about finding out about the effect of drugs?'

I am very fond of Kipling's dictum that, 'a Policy is the blackmail levied on the Fool by the Unforeseen'. It can be dangerous for the applied statistician to sign up wholeheartedly to any particular statistical philosophy. For example, although I do a lot of work that can be classified as Neyman-Pearson I no longer think that the NP approach to bioequivalence is satisfactory (Senn 2001). Of course, neither Neyman nor Pearson ever wrote directly on bioequivalence (as far as I am aware) and what I am criticising is what has been presented to me as the application of the NP method to the bioequivalence problem, or, to follow Deborah Mayo, has been 'dubbed' as NP. But in any case in the end the labels are irrelevant. The real question is what does a given method of analysis deliver? I have a long list of practical problems to test the claims of any method including, the analysis of AB/BA cross-over trials, bioequivalence, sequential trials, multi-centre trials, multiple outcomes, uncontrolled studies, dose-finding studies, two by two tables and linear regression. To me what, is important is how any approach matches up in practice. To take this opportunity to level the record, Senn (2007) is a paper of mine in which I criticise a NP approach to analysing contingency tables. Of course, it may always be claimed that this is not a true NP approach to analysing such tables. But this is exactly the defence that any Bayesian could make (mutatis mutandis) to the examples I gave.

So my article is a reply to the claim that 'Bayes is the only game in town' but there was never any claim of mine that Bayes was a stupid game to play. I do think that the pure form of Bayesianism described by de Finetti is impossible to apply. I do think that the computational game of using vague prior distributions, Bayes theorem and the long-run frequency properties of random numbers to do multiple integration has less to do with this than many suppose. I do not think, however, that intelligent use of Bayesian machinery is something that should be struck out of the applied statistical canon.

So, I am grateful that Deborah Mayo has been inspired to comment on my essay and pleased, of course, that she sees merit in it. However, I am not prepared to go as far as she does. I expect in some contexts so-called Bayesian approaches to be very useful and I expect some frequentist approaches to be problematic. I am not going to throw Bayesian methods out of the statistical toolbox. All statistical approaches have difficulties with nuisance parameters. The subjective Bayesian approach claims to be able to deal with them. The published examples are often unimpressive but this does not mean that frequentist methods deal with nuisance parameters well.

In my world, for all approaches, the devil is in the detail. Deborah Mayo's world is that of working on foundations and I have no quarrel with this, nor do I dispute its importance. As Fisher put it in writing to Maurice Fréchet (8 March 1940), "[o]ur palace of adamant rests upon foundations of gossamer which have to be renewed two or three times a week by the indefatigable labours of mathematical logicians, and yet the superstructure seems to be secure and quite habitable" (Bennett 1990, 129).

References

Bennett, J. H. (1990), *Statistical Inference and Analysis Selected Correspondence of R. A. Fisher*, Oxford: Oxford University Press.

Senn, S. J. (2001), “Statistical Issues in Bioequivalence”, *Statistics in Medicine* 20, 2785–2799.

Senn, S. J. (2007), “Drawbacks to Noninteger Scoring for Ordered Categorical Data”, *Biometrics* 63(1), 296–298; discussion: 298–299.