PROF D. V. LINDLEY
"WOODSTOCK"
QUAY LANE
MINEHEAD
TA24 5QU
TEL: 01643 705 189

Professor Carlos Pereira

Institute of Mathematics & Statistics

University of São Paulo

CP 66281 SAO PAULO 05315-970, BRAZIL

thombayes@aol.com

4 May 2006

Dear Carlos,

A few months back I was sent a paper by you, with Stern and Wechsler, "Can a significance test be genuinely Bayesian?" which, at the time, was read cursorily, found interesting and put aside for careful reading when there was more time. I work very slowly nowadays, so that it is not until now that the paper has been given the attention it deserves and these comments produced. I apologize for the delay and thank you for sending me the paper.

What I like about the paper is its basic use of a concept of evidence for a hypothesis, *ev(H)* in your notation, a concept that does not involve any considerations of decisions or loss functions, yet can be used in decision-making. Put differently, the primary role of a statistician is to make inferences. This was certainly the principle behind the founding of the Royal, and other, Statistical Societies. It is important that the statistician's inference contributions can be used, perhaps by others, in making decisions. Your Table 1 provides a good example: what evidence is there that Ed and Joe are equally exigent, the inference, and then the decision about what to do with the dentists?

Several proposals have been made as to how *ev(H)* might be calculated. Two are p-values and posterior probabilities, to which you add a third. Suppose we look at *ev(H)* as an abstract concept and ask ourselves what properties it should have. For example, if *A* and *B* are two hypotheses, we should be able to order them, to say that *ev(A)* is less or more than *ev(B)*. If this is accepted, then I find it compelling that *ev(A)* should satisfy the assumptions SP_1 to SP_4 in chapter 6 of DeGroot's 1970 book, listed in your admirable list of references. His SP_5 enables evidence to be put on a numerical footing by reference to a standard, the uniform distribution on the unit interval. Accepting all five assumptions, DeGroot proves that *ev(H)* must obey all the rules of probability, in particular that

$$ev(A \text{ or } B) = ev(A) + ev(B) - ev(A \text{ and } B). \tag{1}$$

As you clearly point out in the first paragraph of p.3, your form of *ev(H)* does not satisfy this rule but rather

$$ev(A \text{ or } B) = \max[ev(A), ev(B)]. \tag{2}$$

This implies that your measure of evidence violates at least one of DeGroot's asumptions. Which is it? Are you happy with the violation? The papers of Stern, referred to in the

same paragraph, may address this point but I do not have access to them.

Separate from the axiomatic approach, many people, including myself, have objected to (2), and the concept of possibility, on the grounds that with it $ev(A$ or $B)$ can be calculated from $ev(A)$ and $ev(B)$ without any consideration of the relationship between A and B. This can only be done with probability if A and B are exclusive. Probability requires three numbers adequately to describe the relationship between two hypotheses; possibility uses only two and, for that reason, is often thought to be inadequate. Notice also that possibility lacks an axiomatic basis.

So, I am sorry, I do not support FBST, though you may well have answers to the questions above. If so, I should like to hear them. Your paper does raise some other issues; for example (p.2,l.17) it is not true that the normalized likelihood can replace the posterior because likelihood is not additive, whereas probability is. Bayes factors have similar difficulties; see a paper of mine in *J. Stat. Planning & Inference* **61**, 181–189 (1997). I did not understand what your meant (p.4,l.−15) by "changing the nature of H". It is often forgotten that θ refers to some feature of the real world and dentists (Table 1), like others, have opinions about reality. It is that reality that influences $Pr(H)$. Why (p.7,l.7) are "conclusions based on p−values" subjective? Frequentists often attack Bayesians for being subjective, claiming that $Pr(data \mid \theta)$ and quantities like p−values are objective. Finally, not too seriously, why cite von Foerster's nonsense on p.12, which reads like the pompous rubbish of the structuralists?

You may, after reading the above, think that I have a poor opinion of your paper. Not at all, it is valuable to explore these ideas and to study sharp hypotheses. You have done us all a service and made me think hard and profitably about these ideas, for which I thank you.

How are you these days Carlos? It is a long time since we had such good discussions in São Paulo. Joan and I live very quietly here by the sea and provided all stress is avoided my health is good. I keep up with some statistics and it was a real joy quietly to study your paper. Joan asks me to give her regards to you and your wife, in which I join.

Best wishes,

Dennis