Bayesian Analysis (2006) **1**, Number 3, pp. 451–456

Frequentist Bayes is Objective (Comment on Articles by Berger and by Goldstein)

Larry Wasserman*

Abstract. In this comment, I argue that Bayes procedures with good frequentist properties are objective. I introduce the idea with a short play, followed by some commentary.

Keywords: coverage, frequency probability, objective Bayesian inference

1 Preface: A Short Play

Scene I

Professor van Nostrand, renowned Bayesian statistician, sits in a restaurant in Riomaggore in the Italian Riviera. He enjoys an exquisite lunch and an expensive bottle of wine. A stranger with tattered clothes and a disheveled appearance approaches van Nostrand.

Stranger: At last I have found you!

van Nostrand eyes the stranger wearily.

van Nostrand: Do I know you?

Stranger: I see you don't recognize me. I am Professor H.E. Pennypacker, Nobel prize winner.

van Nostrand: Ahh! Pennypacker. I did not recognize you. Of course, I remember you. It was the huge consulting fees you paid me all those years that made this lavish retirement of mine possible.

Pennypacker: No doubt. That's why I'm here. I've searched high and low for you. I want all that money back you fraud!

*Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, http://www.stat.cmu.edu/~larry

van Nostrand becomes disturbed by Pennypacker's affect. Clearly he has succumbed to dementia.

van Nostrand. Calm down Pennypacker. Have a seat. Perhaps a glass of wine will soothe your mind.

As Pennypacker sits down a waiter brings a glass and pours him some wine.

van Nostrand: Now tell me what this is all about.

Pennypacker: You will recall, my old friend, that you helped me analyze many data sets.

van Nostrand: Of course. I remember each problem quite clearly. And I recall that on each occasion I was quite thorough. I interrogated you in detail, determined your model and prior and produced a coherent 95 percent interval for the quantity of interest.

Pennypacker: Yes indeed. We did this many times and I paid you quite handsomely.

van Nostrand: Well earned money I'd say. And it helped win you that Nobel.

Pennypacker: Well they retracted the Nobel and they took away my retirement savings.

van Nostrand: But why?

Pennypacker: A little thing called scientific objectivity.

van Nostrand: Whatever are you talking about?

Pennypacker: You see, physics has really advanced. All those quantities I estimated have now been measured to great precision. Of those thousands of 95 percent intervals, only 3 percent contained the true values! They concluded I was a fraud.

van Nostrand: Pennypacker you fool. I never said those intervals would contain the truth 95 percent of the time. I guaranteed coherence not coverage!

Pennypacker: A lot of good that did me. I should have gone to that objective Bayesian statistician. At least he cares about the frequentist properties of his procedures.

van Nostrand: Well I'm sorry you feel that way Pennypacker. But I can't be responsible for your incoherent colleagues. I've had enough now. Be on your way.

Pennypacker: Not so fast van Nostrand. Remember this?

Pennypacker pulls out a well worn document.

Pennypacker: This is the consulting contract we signed.

van Nostrand: I remember. The contract stated my fees quite clearly.

Pennypacker: Yes but you apparently never read the fine print. There's a little clause there about scientific objectivity.

Pennypacker smiles and stands up.

Pennypacker: My lawyer will be in touch very soon.

Scene II

Professor Pennypacker, renowned physicist, sits in a restaurant in Riomaggore in the Italian Riviera. He enjoys an exquisite lunch and an expensive bottle of wine. As he lifts his glass, he sees that his wine bottle is empty.

Pennypacker: Waiter! Another bottle. Presto!

The waiter comes running to the table. It is, of course, Dr van Nostrand.

van Nostrand: Yes Professor! Right away.

Pennypacker smiles as van Nostrand scurries to the kitchen.

2 Introduction

The play is supposed to suggest the following thought experiment. You engage in a sequence of unrelated problems and you must choose between two alternatives:

1. A sequence of subjective analyses that yield poor frequency coverage or

2. A sequence of analyses with good frequency coverage that all ignore subjective input.

There are cases where both are simultaneously possible, at least approximately. But I am asking the reader to consider cases where this is not so. Which would you choose? I would choose the latter; I think it would be hard to defend the former.

I interpret "Objective Bayesian inference" to mean "Bayesian methods that have good frequency properties." I refer to these methods as Frequentist-Bayes.

3 Objections to Frequentist-Bayes

Here are some responses I got and my replies.

Frequency probability is not well-defined. This is false. There are at least three good definitions of frequency probability.

The first is physical randomness. Physics provides us with real examples of physical frequency probabilities. For example, in quantum mechanics, if we send a sequence of electrons prepared in the same state through a Stern-Gerlach device set appropriately, 1/2 will emerge spin up and 1/2 will emerge spin down. We can't observe an infinite sequence but we can observe trillions of replications.

The second is by way of algorithmic complexity theory. Here, a sequence is random if it is impossible to substantially compress the sequence. Although early attempts at such definitions by von Mises and Kolmogorov ran into trouble, modern complexity theory provides rigorous definitions. See Li and Vitanyi (1997).

The third is to treat frequency probability by way of idealizations. We take the notion of an unpredictable, infinite sequence with a limiting frequency as a primitive. We then regard random number generators as practical approximations to this Platonic ideal; we do this all the time. A straight line is an idealization; rulers are practical approximations. In subjective probability, a coherent agent with numerically quantifiable beliefs is an idealization; real humans are practical approximations (in the eyes of the subjectivist anyways).

To suggest, as some participants in the conference did, that frequency probability is less well-defined than subjective probability is simply not true.

Coverage is not enough. Some raised the following objection: you can have procedures with good coverage that sometimes yield unacceptable results, such as an empty confidence set. This does not invalidate coverage as a desirable property; it merely shows that coverage is a weak property. The solution is to demand more than just coverage, not to throw out coverage.

For example, in certain low count physics experiments, the standard intervals obtained by inverting Neyman-Pearson tests can be very narrow, or even empty (Mandelkern 2002). This doesn't mean coverage is a bad property; it means we want more from the interval than just coverage. For example, we might also want the interval to represent the uncertainty in the corresponding point estimator. In Wasserman (2002) I proposed to use the narrowest intervals such that (i) they have correct coverage and (ii) their length is always greater than the risk of the estimator. Then the intervals have two properties: they have coverage and they represent estimator uncertainty. Empty intervals are eliminated.

When coverage is not enough, we should ask for more properties in addition to coverage rather than abandoning coverage.

The same is true of subjective Bayesian inference. You can follow the rules of subjective Bayes and still produce unacceptable answers. The response of the subjectivist is that following the rules is only a start: we certainly can distinguish a good subjective analysis from a bad one. The same applies to frequency behavior. Coverage is necessary but not sufficient. I'm not suggesting that we choose any old method with correct coverage and declare victory; I'm suggesting this is a *minimal requirement*.

4 What is Objective Bayes?

I was surprised that Jim Berger stopped short of defining objective Bayes. I have defined a Bayesian procedure to be objective if it yields good frequency behavior. Jim, perhaps more than anyone, has shown time and again how to do this. In some problems, carefully chosen Bayes procedures are easy to implement and they yield good frequency behavior. This is a great success for Bayesian inference. Also, it can be a reference point for a subjective analysis.

5 Subjective versus Elicited

Michael Goldstein made a very important point in his presentation. He does not equate subjective Bayes with elicitation of priors. I think he even remarked that elicitation is a bad idea. I was surprised that this did not raise objections from the participants as I believe that most subjectivists do equate these two things. I don't yet understand the brand of subjectivism that Michael advocates but it does seem deeper and more subtle than the usual "tell me your prior" kind of subjective Bayes.

6 Renormalized Bayes

Is it possible to have subjective input and have good frequency properties? Perhaps. In Genovese et al. (2005) we developed a theory of multiple testing that allows subjective input while preserving frequency properties. As a simple example, consider Bonferroni-adjusted testing of m hypotheses where we reject a hypothesis if the p-value P_i is less than α/m. This controls familywise error at level α. To allow for subjective inputs we consider prior weights $v_1, \ldots, v_m$. Renormalize the weights to $w_i = v_i/\overline{v}$. Then the Bonferroni procedure based on the weighted p-values P_i/w_i still controls familywise error at level α. The same is true for other familywise tests and for false discovery control procedures. I am currently extending this idea to confidence intervals. I refer to methods that modify subjective input to preserve frequency properties as *renormalized Bayes*. These methods, still in their infancy, provide a nice bridge between subjective and objective methods.

References

Genovese, C., Roeder, K., and Wasserman, L. (2005). "False discovery control with p-value weighting." Technical report, Carnegie Mellon University. 456

Mandelkern, M. (2002). "Setting confidence intervals for bounded parameters." *Statistical Science*, 17: 149–172.

Ming, L. and Vitany, P. (1997). *An Introduction to Kolmogorov Complexity and Its Applications*. New York: Springer.

Wasserman, L. (2002). "Comment on Mandelkern." *Statistical Science*, 17: 163.