COMMENTS ON A PAPER BY GARZELLI AND GIUNTI

F. James
CERN ,CH-1211 Geneva 23, Switzerland

Abstract

The paper “Bayesian View of Solar Neutrino Oscillations”, by M. V. Garzelli and C. Giunti, lists eight reasons for using Bayesian methods. I shall examine these reasons critically.

1 The Motivation

In a recent paper [1] by Garzelli and Giunti, eight arguments are presented in favour of using Bayesian methods for statistical analysis of experimental data. I will show that all of these arguments can easily be turned round to give equally convincing arguments *against* the use of Bayesian methods. This is not intended to be a demonstration that there is anything wrong with Bayesian methods; it is only to show that the arguments presented here, which are typical of those being proposed these days by Bayesian extremists, do not stand up to critical analysis.

2 The Sources

The authors refer to several works which have inspired the arguments given in their paper, and to which the reader is directed for further discussion. Some comments are in order concerning these sources, since the authors are all physicists and all have somewhat unusual views about quantum mechanics (QM). I think their views on QM are not unrelated to their Bayesian views. The authors of these works are:

- **Harald Jeffreys.** Although Jeffreys lived through the exciting period of the birth of QM, he did not participate in it. He was active in geophysics, mathematical physics, hydrodynamics and celestial mechanics, none of which need QM.
- **Tom Loredo.** This author makes several statements which make it clear that he agrees with Jaynes that quantum mechanics is a mistake and there is no true randomness in nature. Probability is a property of the state of our knowledge, and not a property of the system under study. See, for example, the footnote on page 92 of [3], which is reference 27 of [1].
- **E. T. Jaynes** made it very clear he did not believe in QM. He even said that physics stopped making progress in 1927 when the world agreed at the Solvay Conference of that year to adopt QM.
- **Giulio D’Agostini**. His views on QM are stated on pp. 285-286 of [4]. I leave it to the reader to decide if he believes in the usual view of QM or not.

I conclude that none of these physicists who are Bayesian authorities believed or believes in the basic proposition that nature is intrinsically random on the level of QM. And furthermore, I believe that this is so because QM is incompatible with the Bayesian concept of probability.

3 The Eight Arguments

Each argument given in [1] is quoted here from the paper (indented), followed by my counter-argument.

3.1 Argument 1

> “All human statements, including scientific ones, represent knowledge (belief) with some degree of uncertainty. Bayesian Probability Theory allows to quantify this uncertainty through the natural definition of probability as degree of belief.”

Bayes Probability Theory (BPT) puts knowledge and belief together in one *pdf*. I would prefer to distinguish between what is known from the experiment and what was believed before it, especially when I read the results of someone else's experiment. BPT does not allow this.

3.2 Argument 2

> "Bayesian Probability Theory allows to calculate through Bayes Theorem the improvement of knowledge as a consequence of experimental measurements. This is how our mind works and how science improves. Therefore, Bayesian Probability Theory is the natural statistical tool for scientists."

Bayes Theorem does not exactly give you the *improvement* of knowledge. It gives only the *improved state* of knowledge provided you input the *prior knowledge*. There is no way to divide out the prior to find out what you learned from the experiment alone.

3.3 Argument 3

> "Probability of an event in Frequentist Statistics is defined as the asymptotic relative frequency of occurrence of the event in a large set of experiments. Obviously such set is never available in practice. Therefore, Frequentist Statistics is based on imaginary data. On the other hand, all inferences in Bayesian Probability Theory are based only on the data that actually occurred."

At least in principle, all frequentist probabilities can be measured to any desired accuracy. That is enough for the theory to be valid. The same thing happens in the definition of the electric field, which involves going to a limit where the test charge is zero, a limit which is physically impossible because charge is quantized. This is not an argument against the existence of electric fields.

Bayesian probabilities on the contrary, are not always measurable, even in principle. Bayesian authority O'Hagan says on page 106 of [2]: "*Any practical usage of Bayesian methods must acknowledge that probabilities cannot be asserted with perfect accuracy.*" In the example given there, the best possible accuracy in P is only 0.25.

3.4 Argument 4

> "Because of the definition of probability in Frequentist Statistics, the results obtained have usually good properties (*i.e.* they are reliable) only in the case of large data sets. In frontier research often a small number of experimental data are available (as in the case of solar neutrinos) and the scientist is not interested in long-term behavior of inferences, but in getting the best possible inference from the available data. Bayesian Probability Theory satisfies this wish."

It is hard to understand what the authors mean here unless you are familiar with several previous papers. The problem they raise is real and has to do with the fact that frequentist statements of probability or confidence are not statements about the particular data obtained, but rather about the ensemble of data that would be obtained if the experiment were repeated. In the rare cases where this produces counterintuitive results, the problem can be solved only by invoking subjective Bayesian methods, and we hesitate to publish subjective measurements of physical constants.

3.5 Argument 5

> "Results in Frequentist Statistics are based on hypothetical data, *i.e.* data that could have been observed but did not occur. Often different scientists may have different ideas on

> which are the relevant hypothetical data, leading to different conclusions (for example, the so-called "optional stopping problem" is well known and discussed in the literature)."

This refers to the Likelihood Principle which forbids Bayesians from making use of the knowledge of the probabilities of observing any data except the data they observed in the actual experiment. So, for instance, if our hypothesis is that the data is Poisson distributed with $\mu = 3.5$ and the data observed are $N = 2$, Bayesians are only allowed to use in their analysis the Poisson probability for $N = 2$, and must pretend they don't know the probabilities for all other values of N. This means that Bayesians are not allowed to use the Chi-square Test and many other important techniques.

3.6 Argument 6

> "In Frequentist Statistics there is no way to take into account theoretical and systematic errors, that are not random variables. Nevertheless, since the great majority of scientific measurements suffer of systematic errors and scientific inferences depend on theoretical errors (they are both crucial in the analysis of solar neutrino data), frequentists treat these errors as if they were random variables. Nobody knows the meaning of results obtained in this way, with an inconsistent method. Bayesian Probability Theory obviously can treat theoretical and systematic errors on the same footing as statistical ones, leading to consistent results."

This is indeed a serious problem, because frequentist statistics is based in random variables, and systematic errors are not (usually) random variables. If the authors had left off the last sentence, this paragraph would be perfectly correct. The problem is that while Bayesian methods are well adapted to many systematic errors (for example, errors in the theory, which really are a matter of belief), they are not well adapted to random errors (where you really want a theory of *random* variables).

3.7 Argument 7

> "Once a problem is well posed the application of Bayesian Probability Theory is clear and straightforward and leads to unique results. On the other hand, in the framework of Frequentist Statistics several arbitrary and difficult choices that can lead to different results are necessary. That is why popular methods (as least-squares) that sometimes have poor performances are widely used without a real understanding of the motivations."

See below.

3.8 Argument 8

> "Since Bayesian Probability Theory allows to calculate the improvement of knowledge as a consequence of physical observations, a prior knowledge is necessary. It has been argued by advocates of Frequentist Statistics that prior knowledge is subjective, leading to undesirable subjectivity in the derivation of scientific results (forgetting the often less clear subjective choices of method, estimator, etc. necessary in Frequentist Statistics). On this problem widely discussed in the literature we want only to remark that all human activities, including scientific research, have some degree of subjectivity, but communication and collaboration among people working in a field allow to reach an agreement on the most reasonable prior knowledge in the field and a way to quantify it. Once the prior knowledge is fixed, Bayesian Probability Theory leads to unique conclusions."

Arguments 7. and 8. together raise two different problems in frequentist analysis:

7.1: *Popular methods (approximations) that sometimes have poor performance.* This is a problem of corrrect use of frequentist methods, not a fundamental problem of the methods themselves. But it is important and means we all have a duty to educate.

7.2: *Arbitrary and difficult choices of method, estimator, etc.*. The Neyman construction does allow the freedom to choose among different confidence intervals, all of which have exact coverage. This is only a second-order arbitrariness in the sense that if one limit is fixed, coverage determines the other limit uniquely.

Note that exactly the same arbitrariness exists in the calculation of Bayesian credible intervals, where Bayes' Theorem yields a *pdf* in the parameters, and it is necessary to choose an interval containing 90% belief. This is actually worse than the frequentist case, because central Bayesian intervals are always two-sided and never give just an upper limit. The usual solution is to take the interval containing the highest values of probability density, but that is metric-dependent, so it is also arbitrary.

In addition to this second-order arbitrariness, Bayesian methods also have a first-order arbitrariness which makes it possible to obtain any limits, both upper and lower, no matter what the data are.

References

[1] *Bayesian View of Solar Neutrino Oscillations*, by M. V. Garzelli and C. Giunti, JHEP 0112 (2001) 017

[2] Anthony O'Hagan, *Kendall's Advanced Theory Of Statistics*, Vol. 2B (1994), Chapter 4.

[3] T. J. Loredo, *From Laplace to Supernova 1987a: Bayesian Inference in Astrophysics*,in *Maximum-Entropy and Bayesian Methods*, Dartmouth, 1989, ed. P. Fougere, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990, pp. 81–142

[4] F. James, L. Lyons and Y. Perrin eds., *Workshop on Confidence Limits* CERN 2000-005