Answers to Fred James' "Comments on a paper by Garzelli and Giunti"

Carlo Giunti
INFN, Sezione di Torino, and Dipartimento di Fisica Teorica, Università di Torino, Via P. Giuria 1, I–10125 Torino, Italy

Abstract
Criticisms of Ref. [1] presented in Fred James' Talk [2] are briefly answered. My point of view on the relationship between Bayesian Probability Theory and Quantum Theory is presented in Section 2..

1. Introduction

First I would like to thank Fred James for reading and commenting on the remarks on Bayesian Probability Theory and Frequentist Statistics presented in Ref. [1]. I would also like to emphasize that those remarks were simply intended to provide a short summary of important issues in Bayesian Probability Theory and Frequentist Statistics that are extensively discussed in several articles and books, such as those in Refs. [3, 4, 5, 6, 7], cited in Ref. [1]. The argument of our paper was the Bayesian analysis of solar neutrino data in terms of neutrino oscillations, not the review of important issues in Bayesian Probability Theory and Frequentist Statistics.

The issues raised by Fred are interesting and I think that it is useful to meditate on them. I think that exhaustive answers to most of his questions and criticisms can be found in the existing literature (see Refs. [3, 4, 5, 6, 7]). In the following limited space I will try to clarify in a concise way the issues raised by Fred examining one by one his points in Ref. [2]. Most of my remarks follow from my readings and are, therefore, not original. An exception, as far as I am aware, is my point of view on the very important problem of the relationship between Bayesian Probability Theory and Quantum Theory, presented in Section 2..

2. The Sources: Bayesian Probability Theory vs. Quantum Theory

Like Fred, many people think that Bayesian Probability Theory and Quantum Theory are incompatible. This apparent incompatibility stems from the different definitions of "probability" in Bayesian Probability Theory and Quantum Theory.

In Bayesian Theory, the "probability" of something is a quantification of the subjective degree of belief in that something, that I will call *Bayesian Probability*. It is assumed that this degree of belief follows from a rational evaluation of all the available information. Two individuals with the same information should hold the same Bayesian Probability. Thus, Bayesian Probability can be considered as a quantification of the degree of knowledge reached by individuals or groups of individuals sharing the same information.

In Frequentist Statistics, the "probability" of an event, that I will call *Frequentist Probability*, is defined as the relative frequency of occurrence of the event in an infinite set of experiments.

In Quantum Theory[1] the squared modulus of an amplitude, called "probability", gives the asymptotic relative frequency of occurrence of some event. This "probability", that I will call *Quantum Probability*, is a Frequentist Probability.

[1] I strongly disagree with the criticism to Quantum Theory in Chapter 10 of Jaynes' book [6], an otherwise enlightening book, in my opinion. I think that Jaynes' claim that "physical probabilities" do not exist is unfounded and unnecessary. I am astonished by his claim that physicists do not think seriously about the problem of the probabilistic interpretation of Quantum Theory (never heard of the famous debate between Einstein and Bohr? of the Einstein-Podolsky-Rosen paradox? of Bohm's hidden variable theory? of Bell inequalities? of recent experiments confirming the violation of Bell inequalities? etc...). I think that laughing is the best answer to his statement that no progress in physics (except for the discovery of new particles and calculation techniques) has been made since 1927.

Now, the questions are: Is there a contradiction between Quantum Probability and Bayesian Probability? Can the two live together?

I think that there is no contradiction between Quantum Probability and Bayesian Probability and the two can happily live together, as long as the scientist is conscious of their difference and knows which one she/he is using.

The point is that Quantum Theory and Bayesian Probability Theory are used for different tasks. Quantum Probability gives the theoretical asymptotic relative frequency of occurrence of some event, while Bayesian Probability quantifies our degree of knowledge following, for example, the analysis of experimental data.

One can happily use Quantum Theory to predict relative asymptotic frequencies and Bayesian Probability Theory to quantify one's degree of knowledge. Furthermore, believers of Quantum Theory can consciously transform Quantum Probability to Bayesian Probability with the same value (not vice-versa!).

Indeed if there is a theory in which one believes which provides the values of relative frequencies, it would be foolish to assume different values for the subjective probabilities. A trivial example: in the toss of a fair coin the asymptotic relative frequency of heads is 1/2 and it would be utterly foolish to assign to heads a Bayesian Probability different from 1/2.

The transformation of Quantum Probability into Bayesian Probability is often necessary in the application of Bayes' Theorem

$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)}\,, \tag{1}$$

where H represent hypotheses and D represent data. In many experiments the sampling probability $P(D|H)$ is calculated using Quantum Theory and the Quantum Probability is correctly transformed to Bayesian Probability with the same value.

3. The Eight Arguments

3.1 Argument 1: Knowledge and Belief

I guess that the distinction between knowledge and belief may make the happiness (and career) of a philosopher. As a simple scientist I know what I believe and I believe what I know to be likely to be true. I could also say that I do not believe what I do not know and I do not know most of the things in which I do not believe.

If Fred does not believe what he knows to be likely true he is free to make two analyses: one with probability = degree of belief and one with probability = degree of knowledge. Why does Bayesian Probability Theory not allow this?

Jokes apart, Jaynes' book [6] clearly explains that Bayesian probability should be evaluated taking into account all the available information and two persons with the same information should evaluate the same probability. Since both knowledge and belief follow from information, it seems to me that they are tightly connected in a rational mind.

3.2 Argument 2: "improvement" or "improved"?

Fred is right, but I think that it is rather obvious that we meant "improved". Sorry for the misprint!

3.3 Argument 3: Measurability and Authorities

Fred seems to judge Bayesian Probability Theory with Frequentist criteria. Measurability of probability is an absurd idea in Bayesian Theory. What does it mean to measure one's degree of belief? What is its usefulness?

I do not understand much of what O'Hagan writes in paragraph 4.26 on page 106 of Ref. [8] and I think that one should not just follow "authorities" but Reason. Blindly following authority is the end of science. For example, I cited Refs. [3, 4, 5, 6, 7] not because the authors are authorities, but because I like most of what they have written there.

¿From what I understand of what O'Hagan writes in paragraph 4.26 on page 106 of Ref. [8], I have the impression that he is confusing Bayesian probability as degree of belief with Frequentist probability as relative frequency of events. In particular, the sentence *"Any practical usage of Bayesian methods must acknowledge that probabilities cannot be asserted with perfect accuracy"* is for me nonsense.

In Frequentist Statistics probability is considered to be an objective property of some apparatus and I agree that it can be measured with any desired accuracy. However, I guess that Fred will have to agree that, as all measurable objective properties, Frequentist probability is an unknown quantity, or, using the words of O'Hagan, Frequentist "probabilities cannot be asserted with perfect accuracy".

Therefore, Bayesian probability is known but not measurable, and Frequentist probability is measurable but unknown. You chose which you prefer.

The trivial example of tossing a fair coin can clarify the above remarks about Frequentist probability. Even after billions of tosses the relative frequency of heads is probably different from exactly 1/2. How many tosses are necessary to say that the probability of heads is exactly 1/2? When can one stop?

In spite of the fact that the probability to get heads in the toss of a fair coin is never measured with perfect accuracy[2], I guess that even orthodox Frequentists take it to be exactly 1/2.

It is a fact of life that the values of Frequentist probabilities are not obtained from measurement, because the definition of Frequentist probability requires "unrealizable experiments", as noted in Ref. [9], Section 2.1.1. Instead, Frequentist probabilities are obtained by abstract reasoning, or by theory (as, for example, Quantum Mechanics).

3.4 Argument 4: Best Inference

See, for example, Ref. [10]. I emphasize again that our paper in Ref. [1] could not and was not intended to provide an exhaustive presentation of Frequentist Statistics and Bayesian Probability Theory. It is quite obvious that a beginner should consult the vast literature on these subjects in order to gain some understanding.

3.5 Argument 5: Hypothetical Data

Before this Conference I did not know the "Likelihood Principle". After hearing about it I am not impressed.

The point, that I think we wrote rather clearly, is that if your conclusions are based on hypothetical data, they depend on which hypothetical data are in your mind.

The fact that Bayesian Probability Theory allows conclusions to be drawn without need of hypothetical data seems to me a desirable feature. As far as I know, in Bayesian Probability Theory the use of hypothetical data is not forbidden, it is just not needed. I guess that also Frequentists (as any reasonable person) would agree on the undesirability of using hypothetical data if they did not need them desperately.

Frequentists like very much to teach what is allowed and what is forbidden in Frequentist Statistics, whereas, in my experience, Bayesians are much less dogmatic. People are not allowed or forbidden to do something, but they are expected to understand what makes sense and what does not, what is useful and what is not, etc...

[2]Somebody may say that I am writing nonsense, because a fair coin has by definition probability 1/2 to give head when tossed. Then, I ask: how do I know that I have a fair coin? It is clear that we return to the measurement problem.

3.6 Argument 6: Systematic and Statistical Errors

I am glad that we agree on something.

What is wrong with the last sentence: "Bayesian Probability Theory obviously can treat theoretical and systematic errors on the same footing as statistical ones, leading to consistent results."?

Of course one can think that "Bayesian methods ... are not well adapted to random errors", but this has nothing to do with the fact that "Bayesian Probability Theory obviously can treat theoretical and systematic errors on the same footing as statistical ones, leading to consistent results."

3.7 Argument 7: Which Approximate Method?

It is true that the problem of choice of an appropriate approximate method is not a fundamental problem, but it seems to me a serious problem in real life. In spite of the fact that most physicists are educated to think as Frequentists, very few of them understand the subtleties of Frequentist Statistics and possess knowledge beyond the simplest methods. I do not think that anybody can say that this is due to the fact that physicists are stupid or lazy to learn. It is due to the vastness and complexity of the theory of Frequentist Statistics. In order to master the theory of Frequentist Statistics a physicist should devote a large fraction of his time to the study of Statistics, with his work in physics suffering. On the other hand, the principles of Bayesian Probability Theory are simple, easy to master and allow the scientist to make useful inferences without much further learning.

3.8 Argument 8: Arbitrariness

I have the impression that again Fred is applying Frequentist criteria to Bayesian methods. It is well known that in Frequentist Statistics one must choose the method independently from the data in order to have coverage (see, for example, Refs. [9, 11, 12, 13]). A popular method is that of central intervals, which has sometimes desirable properties (see [11, 14]). On the other hand, in Bayesian Probability Theory one can choose the interval that better represents the posterior p.d.f. after the calculation of the posterior. It would be a clear nonsense to choose a central interval if the posterior is better represented by an upper limit.

For example, Bayesian Probability Theory has often been used to obtain meaningful upper limits on the neutrino mass in Tritium β-decay experiments. As far as I know, nobody has ever had the insane idea to give a central interval when the data did not show any evidence of a finite neutrino mass.

Of course, in spite of reason, one may decide to give a central interval in any case. Then, it is true that in Bayesian Probability Theory one cannot get an upper limit. But in Frequentist Statistics one gets sometimes a central interval, sometimes an upper limit and sometimes an empty interval (nothing, no information). You choose which you think is better.

Speaking about more serious problems, the arbitrariness in Frequentist Statistics and Bayesian Theory arises at different stages of the analysis. As I wrote above, in Frequentist Statistics one must choose the method independently from the data in order to have coverage. The results of an analysis of experimental data is an interval (or a set of intervals for different Confidence Levels). Once these intervals are obtained one cannot decide that the method was inappropriate and switch to another method. However, someone else could prefer another method, independently from the results. The application of the alternative method requires a repetition of the whole analysis starting from the raw data.

In Bayesian Theory the result of an analysis is a posterior p.d.f. that can be published to allow everybody to extract their own interval. Since the posterior p.d.f. depends on the prior, it is highly desirable that experimenters publish also the Likelihood function, which allows everybody to calculate a posterior without the need to re-analyze the raw data (which is a hard job that usually can be done properly only by the experimenters).

4. Conclusions

In conclusion, I would like to thank the organizers of this very interesting Conference for giving me the opportunity to present these answers.

References

[1] M. V. Garzelli and C. Giunti, JHEP **12**, 017 (2001), hep-ph/0108191.

[2] F. James, *Comments on a paper by Garzelli and Giunti*, these proceedings.

[3] H. Jeffreys, *Theory of Probability* (Oxford University Press, New York, USA, 1961), First published in 1939.

[4] T. J. Loredo, (1990), in Maximum-Entropy and Bayesian Methods, Dartmouth, 1989, ed. P. Fougere, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990, pp. 81–142, http://astrosun.tn.cornell.edu/staff/loredo/bayes/tjl.html.

[5] T. J. Loredo, (1992), in Statistical Challenges in Modern Astronomy, ed. E.D. Feigelson and G.J. Babu, Springer-Verlag, New York, 1992, pp. 275–297, http://astrosun.tn.cornell.edu/staff/loredo/bayes/tjl.html.

[6] E. T. Jaynes, *Probability Theory: The Logic of Science* (Fragmentary Edition, June 1994), http://bayes.wustl.edu/etj/prob.html.

[7] G. D'Agostini, CERN Yellow Report **99-03** (1999).

[8] A. O'Hagan, *Kendall's Advanced Theory of Statistics*, Vol. 2B, *Bayesian Inference* First Edition, Halsted Press, 1994.

[9] W. Eadie, D. Drijard, F. James, M. Roos, and B. Sadoulet, *Statistical Methods in Experimental Physics* (North Holland, Amsterdam, 1971).

[10] P. Clifford, CERN Yellow Report **CERN 2000-005**, 157 (2000), Proceedings of the Workshop on Confidence Limits at CERN, 17-18 January 2000, edited by F. James, L. Lyons, Y. Perrin. http://preprints.cern.ch/cernrep/2000/2000-005/2000-005.html.

[11] A. Stuart, J.K. Ord and S. Arnold, *Kendall's Advanced Theory of Statistics*, Vol. 2A, *Classical inference and the linear model* Sixth Edition, Oxford University Press, 1999.

[12] R.D. Cousins, Am. J. Phys. **63**, 398 (1995).

[13] C. Giunti, Phys. Rev. **D59**, 113009 (1999), hep-ex/9901015.

[14] C. Giunti and M. Laveder, Nucl. Instrum. Meth. **A480**, 763 (2002), hep-ex/0011069.