Bayesian Analysis (2010) **5**, Number 2, pp. 223–228

# Comment on Article by Robert

Larry Wasserman*

**Abstract.** This article is a response to Christian Robert's review of *The Search for Certainty* by Krzysztof Burdzy. I provide my own review, some comments on Robert's review and a few general comments about the foundations of probability.

**Keywords:** foundations of probability, frequentist probability, subjective probability

## 1 Introduction

I first found out about Krzysztof Burdzy's book *The Search for Certainty* when I read reviews of the book on the blogs of Christian Robert and Andrew Gelman. Both reviews were negative so I was inclined to ignore the book. However, curiosity — and the convenience of one-click ordering — got the best of me, and I ended up buying and reading the book. While the book isn't perfect, I thought the reviews on the blogs were far too harsh. I decided to write my own review which I sent to my friends Andrew and Christian. Andrew kindly posted my review on his blog.

When Christian's review was accepted for this journal the Editor asked if I would submit a discussion based on my own review. What follows is an adapted version of my original review. I will begin with some general comments, followed by my review of the book, followed by my review of Christian's review and ending with a few comments on the foundations of probability.

## 2 Some General Remarks

Burdzy is a probabilist from the Department of Mathematics at the University of Washington. But his book is about the philosophical foundations of probability, not about the mathematical aspects of probability. He is addressing a very old question that we have all grappled with: what does it mean to assign a probability to an event? For proving mathematical theorems about probability, this question is uninteresting. But for statisticians, the question is not pedantic: it gets to the very heart of what we do every day.

If you ask ten statisticians what they mean when they say "the probability of heads is 1/2" you will get ten different answers. Broadly speaking, most statisticians would declare their philosophy to be either subjectivist or frequentist. It is thus fair to ask: what author has provided a sound and influential foundation for your school of thought?

For subjectivists, a likely answer is de Finetti. For frequentists, the answer is less

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

 DOI:10.1214/10-BA601A

clear. Burdzy identifies Richard von Mises as the best possibility. This is debatable but I cannot think of a better candidate. One minor complaint about choosing von Mises is that von Mises collectives were long ago abandoned in favor of random sequences as defined in the modern theory of computational complexity. But let us stipulate for the purposes of this discussion that de Finetti and von Mises can serve as icons for the two schools of thought.

Burdzy considers both to be failures. The main reason is that neither de Finetti nor von Mises (nor the computational complexity replacement of von Mises) tells us how to actually assign a probability to a single event.

De Finetti tells us that probability is subjective but does not say how we should come up with the numbers. Most Bayesians believe that one should base subjective probabilities on personal prior experience. But in addition to being vague, that prescription is not really part of de Finetti's theory.

I have to admit my bias here. I think science should be as objective as possible. Subjective probability is thus unappealing to me. Some say that some subjectivity is unavoidable. That's true but that's not a reason for making subjectivity the basis of our field. It's a bit like saying that disease is inevitable so let's embrace disease and celebrate it. I'd rather try to eliminate it. Anyway, my bias against subjective probability inevitably leads me to be more receptive to Burdzy's criticisms of de Finetti. Also, I read de Finetti. I wonder how many Bayesians have. I defy anyone to read de Finetti and not be left with a sense that he was kidding, or very disturbed.

In practice, statisticians typically ignore (or misrepresent) the philosophical foundations espoused by de Finetti (subjectivism) and von Mises (frequentism). This is itself a damning criticism of the supposed foundational edifice of statistics.

# 3 Comments on Burdzy

Burdzy makes a convincing case that the philosophy of probability is a complete failure.

He criticizes von Mises because his theory, based on defining limits of sequences (or collectives) does not assign a probability to a given event. There are also technical issues with the mathematical definition of a collective that von Mises was unable to resolve but these can be fixed rigorously using modern computational complexity theory. But that doesn't blunt the force of Burdzy's main criticism. I think the problems with the foundations of frequentist probability rest more with a lack of interest in developing these foundations than with the theory itself. I will discuss this point later.

The problems with subjectivism go deeper in my opinion and, correspondingly, Burdzy's criticism of de Finetti is more thorough. There is the usual criticism, namely, that subjective probability is unscientific as it is not falsifiable. Moreover, there is no guidance on how to actually set probabilities. Nor is there anything in de Finetti to suggest that probabilities should be based on informed prior opinion, as many Bayesians would argue. More surprising is Burdzy's claim that subjective probability has the same problem as von Mises' frequency theory: it does not provide probability for an individual event. This claim will raise the hackles of die-hard Bayesians. But he is right: de Finetti's coherence argument requires that you bet on several events. The rules of probability arise from the demand that you avoid a sure losing bet (a Dutch book) on the collection of bets. The argument does not work if we supply a probability only on a single event. The criticisms of de Finetti's subjectivism go beyond this and I will not attempt to summarize them.

Burdzy provides his own foundation for probability. His idea is that probability should be a science, not a philosophy, and that, as such, it should be falsifiable. Allow me to make an analogy. Open any elementary book on quantum mechanics and you will find a set of laws. These laws can be used to make very specific predictions. If the predictions are wrong, (and they never have been), then the laws would be rejected. But to use the laws, one must inject some specifics. In particular, one must supply the Hamiltonian for the problem. If the resulting predictions fail to agree with reality, we can reject that Hamiltonian.

To make probability scientific, Burdzy proposes laws that lead to certain predictions that are vulnerable to falsification. More importantly, the specific probability assignments we make are open to being falsified. Before stating his laws, let me emphasize a crucial aspect of Burdzy's approach. Probability, he claims, is the search for certainty; hence the title of the book. That might seem counter to how we think of probability but I think his idea is correct. In frequentist theory, we make deterministic predictions about limits of sequences. In subjectivist theory, we make the deterministic claim that if we assign probabilities consistent with the rules of probability then we are certain to be immune to a Dutch book. A philosophy of probability, according to Burdzy, is the search for what claims we can make for certain.

Burdzy's proposal is to have laws – not axioms – of probability. Axioms, he points, merely encode facts we regard as uncontroversial. Laws instead, are proposals for a scientific theory that are open to falsification. Here are his five proposed laws (paraphrased):

(L1) Probabilities are numbers between 0 and 1.

(L2) If A and B are disjoint then P(A or B) = P(A) + P(B).

(L3) If A and B are physically independent then they are mathematically independent meaning that P(A and B) = P(A)P(B).

(L4) If there exists a symmetry on the space of possible outcomes which maps an event A onto an event B then P(A)=P(B).

(L5) P(A)=0 if and only if A cannot occur. P(A)=1 if and only if it must occur.

Some comments are in order. (L1) and (L2) are standard. (L4) refers to ideas like independent and identically distributed sequences, or exchangeability. It is not an appeal to the principle of indifference. Quite the opposite. Burdzy argues that introducing symmetry requires information, not lack of information.

(L3) and (L4) are taught in every probability course as add-ons. But in fact they are central to how we actually construct probabilities in practice. The author asks: Why treat them as follow-up ideas? They are so central to how we use probability that we should elevate them to the status of fundamental laws.

(L5) is what makes the theory testable. Here is how it works. Based on our probability assignments together with the rules of probability, we can construct events A that have probability very close to 0 or 1. For example, A could be the event that the proportion of heads in many tosses is within .00001 of 1/2. If this doesn't happen, then we have falsified the probability assignment. Of course P(A) will rarely be exactly 0 or 1, rather, it will be close to 0 or 1. But this is precisely what happens in all sciences. We can test prediction of general relativity or quantum mechanics to a level of essential certainty, but never exact certainty. Thus Burdzy's approach puts probability on the same level as other scientific theories.

To summarize, Burdzy's approach is to treat probability as a scientific theory. It has rules for making probability assignments and the resulting probabilities can be falsified. Not only is this simple, it is devoid of the murkiness of subjectivism and the weakness of von Mises' frequentism. And, perhaps most importantly, it reflects how we use probability. It also happens to be easy to teach. My only criticism is that I think the implications of (L1)-(L5) could be fleshed out in more detail. It seems to me that they work well for providing a foundation for testable frequency probability. That is, they provide a convincing link between probability and frequency. But that could reflect my own bias towards frequency probability.

My short summary of this book does not do justice to the author's arguments. In particular, there is much more to his critique of subjective probability than I have presented in this review. The best thing about this book is that it will offend and annoy both frequentists and subjectivists. I implore my friends on both sides of the philosophical divide to read the book with an open mind.

## 4 Comments on Robert's Review

Christian's review is very harsh. I can definitely see why someone might not like Burdzy's book but I was surprised at how negative Christian's review is. At the risk of sounding ironic, assessing a books with a philosophical flavor is a very subjective task. So it is not surprising that some people will not like the book and others will. (I notice from Christian's blog that he likes fantasy novels. I hate them. But I think we have similar taste in wine.)

One of Christian's criticisms of the book is a lack of philosophical context. That is, there are few connections to previous work of philosophers on this topic. As he says: " ... the lack of connections with the existing literature on the philosophy of science and epistemology appears to me as an indicator of the isolated position of the author ..." I agree with Christian about the lack of connection with the philosophical literature. But I don't have a problem with this. I don't think there is much to be found in the philosophy literature that is going to interest the intended readers: probabilists and statisticians. Indeed, that's one of the points of the book.

Christian is also disappointed that the author does not follow up his laws with detailed examples. One could cite Harold Jeffreys as an example of this. Again, I think Christian has a point here; more examples would have been better. But I think this criticism is overstated. Burdzy claims he is merely encoding what is already in every probability book. In most texts, the meaning of probability is barely discussed. The usual axioms are given then most of the book is about how to actually construct sophisticated probability models. The latter, he argues, involves deriving models from (L3) and (L4) in one way or another. So why not include them in our axioms? I am not completely convinced about this but I think it is a reasonable position.

Christian goes on to say "It is hard to come up with satisfying epistemological justifications for introducing in 2009 a new set of axioms ... and not to find difficulties with them." I think Burdzy is saying that we all use these axioms (or laws) so why don't we make that formal? He's not really proposing anything new, he is just encoding what we do when we use probability. (I do have one issue with the five laws: I think (L4) is a bit vague and could be stated in more detail.)

Next Christian argues that the book has no relevance for statisticians. To be honest, I think that, in general, the foundations of probability inevitably have limited relevance for statistics. A small fraction of statisticians are interested in foundations and we will read and debate these things. Most statisticians just compute things and don't worry too much about what it all means. And there is nothing wrong with this.

Christian criticizes Burdzy for having too narrow a focus. Indeed, if he had written a longer book, Burdzy could have carried out detailed critiques of Savage, Jeffreys, Wald and so on. I suppose I would have appreciated a broader coverage as well, but a shorter more focused book has some advantages. Christian is also unhappy with the "predominantly non-technical" style of the book but I did not find that to be an impediment; the author was able to communicate his ideas without the need for technicalities which is fine with me. Finally, Christian doesn't like the book's cover! I have to assume that on this point he is joking. Although, a purple cover would have been nice.

## 5 The Foundations of Probability

It is impossible to write a book on the foundations of probability that everyone will like. This book is no exception. I liked it, Christian didn't. So it goes. I'd like to conclude with a few thoughts of my own on the meaning of probability.

I am comfortable with frequency probability. The idea of a long, unpredictable sequence is an idealization much like the idea of a perfectly straight line in geometry. These basic idealizations can be made formal mathematically, physically and pragmatically. Here is a table that makes the point:

| | Straight Line | Random Sequence |
|---|---|---|
| Mathematical | shortest path between two points | defined by computational complexity |
| Physical | path of light in a vacuum | output of sequence of quantum experiments |
| Pragmatic | ruler | random number generator |

The fact that this does not provide a probability statement about a single event does not trouble me. I think of statistics as providing procedures with guarantees that hold when used by an ensemble of users. So probabilities on individual events do not seem central to me.

# 6 Conclusion

Statisticians and probabilists have never agreed on the meaning of probability and probably never will. Burdzy's book will not create a consensus on the meaning of probability nor will everyone agree with his critiques of the most popular interpretations. But it does raise good questions and I recommend the book to anyone interested in the foundations of probability.

**Acknowledgments**

The author wish to thank Cosma Shalizi for helpful comments.