View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by PhilSci Archive

THE DOOMSDAY ARGUMENT WITHOUT KNOWLEDGE OF BIRTH RANK

by Bradley Monton[1]
Department of Philosophy, University of Kentucky
bmonton@uky.edu

September 13, 2001

[1] I thank Nick Bostrom for helpful discussion.

THE DOOMSDAY ARGUMENT WITHOUT KNOWLEDGE OF BIRTH RANK

1. The Carter-Leslie Doomsday argument (Leslie 1989, 1996), as standardly presented, relies on the assumption that you have knowledge of your approximate birth rank. I will demonstrate that the Doomsday argument can still be given in a situation where you have no knowledge of your birth rank. As I will show, this allows one to reply to Bostrom's (2001, 2002) defense of the Doomsday argument against the refutation suggested by Dieks (1992), and independently developed by Kopf et. al. (1994) and Bartha and Hitchcock (1999).

The Doomsday argument runs as follows (following Bostrom 2001). Suppose that you have narrowed the possibilities for doom down to two:

H_1: "there will have been a total of 200 billion humans."

H_2: "there will have been a total of 200 trillion humans."

Let us suppose that these hypotheses agree on the number of humans that exist on Earth from 20:41 to 20:42 GMT on August 15, 2001. This supposition is not a standard part of the Doomsday argument, but it does not affect the Doomsday argument, and it is needed for my argument below. There are two reasons this supposition is reasonable. First, if the hypotheses disagreed on the number of humans that exist during that time period, then in principle it would be easy to falsify one of them, by checking population figures. (This point is made by Dieks 2001, 3.) Second, since the hypotheses are meant to represent the possibilities that doom will come soon and that doom will come late, the hypotheses should be understood as agreeing on the number of humans that exist up to now and into the short-term future; they disagree only about how many humans will exist in the long-term future.

After considering the various ways in which human life might end, you assign the following probabilities:

$Pr(H_1) = 0.05$

$Pr(H_2) = 0.95.$

You also know proposition R: "I am the 60 billionth human to have been born". Reasoning with the Self-Sampling Assumption:

(SSA) Observers should reason as if they were a random sample from the set of all observers in their reference class,

you have the following conditional probabilities:

$Pr(R \mid H_1) = 1/200$ billions

$Pr(R \mid H_2) = 1/200$ trillions.

Bayes' theorem then gives the result that $Pr(H_1 \mid R) = 0.98$. Since you know R, your posterior probability for H_1 is 0.98 – doom is likely to come soon.

2. Suppose that you have no knowledge of your birth rank. How could the Doomsday argument still be given? What is needed is a property p such you know you have p, and the total number having p would be the same regardless of whether H_1 or H_2 is true. We each possess such properties, and thus the Doomsday argument does apply. For me, one such property would be the property of being alone in 1423 Patterson Office Tower in Lexington, Kentucky, from 20:41 to 20:42 GMT on August 15, 2001. Call that property t, and let T be the proposition that someone has property t. Before 20:42 I did not know that T is true, but now I do. I can model this learning that T by conditionalization using my prior probability function Pr^*: for any proposition A,

$$Pr(A) = Pr^*(A \mid T)$$

Note that it is reasonable for Pr^* to be such that the probability of T does not depend on whether H_1 or H_2 is true:

$$Pr^*(T \mid H_1) = Pr^*(T \mid H_2) = Pr^*(T)$$

If this were not the case, then conditionalization on T would shift my probabilities for H_1 and H_2. The reason it is reasonable for Pr^* to be such that T does not depend on H_1 or H_2 is that H_1 and H_2 agree on the number of humans existing on Earth from 20:41 to 20:42 GMT on August 15, 2001. It follows that

$$Pr(H_1) = Pr^*(H_1) \text{ and } Pr(H_2) = Pr^*(H_2).$$

Now, let M be the proposition that I have property t. Reasoning using the SSA,

$$Pr(M \mid H_1) = 1/200 \text{ billions}$$

$$Pr(M \mid H_2) = 1/200 \text{ trillions.}$$

Bayes' theorem then gives the result that $Pr(H_1 \mid M) = 0.98$. Since I know M, my posterior probability for H_1 is again 0.98.

3. The reply to the Doomsday argument given by for example Bartha and Hitchcock (1999) relies on what Bostrom (2001, 382) calls the Self-Indication Assumption (SIA): roughly, "finding that you exist gives you reason to think that there are many observers". The idea behind Bartha and Hitchcock's reply is that conditionalizing on your existence shifts probabilities in favor of H_2, and the Doomsday argument shifts probabilities in favor of H_1, and these two shifts cancel each other out. Bostrom has recently argued against this reply to the Doomsday argument by presenting a scenario for which he claims that the SIA leads to unintuitive results. I will defend the SIA and Bartha and Hitchcock's reply.

Bostrom's (2001, 383; 2002, Chapter 7) scenario is as follows. It is the year 2100, and physicists assign probability 0.5 each to theories T_1 and T_2. T_1 entails that there are a total of a trillion trillion observers, while T_2 entails that there are a total of a trillion trillion trillion observers. We do not know our birth ranks, even approximately. Physicists are going to do an experiment to decide between T_1 and T_2, but before they do a presumptuous philosopher explains that there is no need for the physicists to do the experiment. The presumptuous philosopher says that since he exists, that makes it more likely that there are more observers – T_2 is a trillion times more likely than T_1.

Bostrom's idea here is that, since we have no knowledge of our birth ranks in this scenario, we can only get the first probability shift via the SIA in favor of more observers; we cannot get the second Doomsday shift in favor of fewer observers. But as I have shown, the Doomsday argument can be given even when we have no knowledge of our birth rank. We would have to specify that T_1 and T_2 agree on the number of observers existing in some appropriate spacetime region, but this is a legitimate assumption to make. (We can pick the region such that, if the hypotheses disagreed, then in principle it would be easy to falsify one of them, by checking population figures.) Thus, Bostrom's scenario does not show the unreasonableness of the SIA, and Bartha and Hitchcock's reply to the Doomsday argument is unrefuted.

References

Bartha, P. and Hitchcock, C. 1999: "No One Knows the Date or the Hour: An Unorthodox Application of Rev. Bayes's Theorem". *Philosophy of Science,* 66 (Proceedings), pp. S339-S353.

Bostrom, N. 2001: "The Doomsday Argument, Adam & Eve, UN++, and Quantum Joe", *Synthese* 127: 359-387.

Bostrom, N. 2002: *Anthropic Bias: Observational Selection Effects in Science and Philosophy*, New York: Routledge, forthcoming.

Dieks, D. 1992: "Doomsday – Or: The Dangers of Statistics". *Philosophical Quarterly* 42: 78-84.

Dieks, D. 2001: "The Probability of Doom". Preprint 247, available at http://philsci-archive.pitt.edu.

Kopf, T., Krtous, P., and Page, D. 1994: "Too Soon for Doom Gloom?". Preprint gr-qc9407002, available at http://xxx.lanl.gov.

Leslie, J. 1989: "Risking the World's End", *Bulletin of the Canadian Nuclear Society* May: 10-15.

Leslie, J. 1996: *The End of the World: The Science and Ethics of Human Extinction*. New York: Routledge.