The Probability of Doom

Dennis Dieks
History and Foundations of Science, Utrecht University
P.O.Box 80.000, 3508 TA Utrecht

April 24, 2001

Abstract

According to the Doomsday Argument the probability of an impending extinction of mankind is much higher than we think. The adduced reason is that in our assignment of probabilities to soon or not so soon doom we have not fully taken into account that we live in the specific year 2001. This is relevant information, because if I consider myself as an arbitrary member of the human race I have a much higher probability of finding myself living in 2001 on the hypothesis of a soon extinction, Doom Soon, than on the hypothesis of Doom Late—according to the latter hypothesis there are so many more years I could have found myself living in. Accordingly, Bayesian reasoning leads to a posterior probability of the Doom Soon hypothesis, after I have taken the evidence of my birth date fully into account, that is much higher than the prior probability.

I show that the Argument is nothing but a rather trivial mathematical exercise in the calculation of posterior from prior probabilities; it is only about the *relation* between these probabilities and is silent about the concrete values these probabilities should have. Nothing in the Argument supports the conclusion its proponents think it supports, namely that Doom Soon is much more probable than we ordinarily think. The Argument is formally valid, but ineffective.

1 The Doomsday Argument

There are dangers threatening the future of mankind. Using information about the present political situation in the world, the risks of nuclear technology, the possibility of epidemics, and so on, or perhaps relying more on our intuition, we may attempt to become precise about how serious we think these dangers are. We can do so by assigning a probability to the hypothesis that mankind will become extinct soon, within the present century, say. There will also be a remaining probability, of the possibility that the doom scenario will not turn into a reality. Let us for simplicity assume that there is only one alternative hypothesis: that the human civilisation will flourish for a very long time, with population figures that grow and grow as time goes on. Perhaps the human race will in this case colonize other planets, or even galaxies, to find room for all humans alive in the future.[1]

Suppose I have determined probabilities that measure my degrees of belief in these two hypotheses. Then the following thought occurs to me. In the evaluation of my probabilities I have not completely taken into account that I live now, in the year 2001 AD. Of course, I have used information about the state of the world in 2001; but I have not paid attention to the fact that I myself happen to live in 2001. This is, however, relevant additional information. For according to the Doom Soon hypothesis, that I live in 2001 means that I am a fairly typical human being; after all, a significant proportion of all humans ever alive will live in this century if the Doom Soon scenario corresponds to the actual fate of the world. By contrast, the number of people living in the far future will be overwhelmingly great if the Doom Late hypothesis is true. So on the latter hypothesis the situation in which I actually find myself, living so early in the history of mankind, is a very unlikely one. As a general methodological principle, I should prefer the hypothesis that endows the available evidence with the highest probability. Taking into account that I live now, in the year 2001 AD, should therefore move my odds appreciably in favour of the Doom Soon scenario.

[1]Of course, assuming that there are only two scenarios is unrealistic. A more natural assumption would be that there are many hypotheses to consider, each one with its own prediction for when mankind will die out. However, we could assemble these hypotheses in two groups: those predicting doom before a certain date, and those predicting doom later. We would thus effectively end up with two possibilities, and the argument in the text would apply. Therefore, nothing essential hinges on the restriction to two rival scenarios.

This is the Doomsday Argument, in the form defended by John Leslie. It has been around in the literature for more than ten years ([1, 2]) and continues to be discussed as a controversial, but unrefuted line of reasoning ([4]; see [5, 6] for a survey of the extensive literature that has appeared on the Argument). In his book, *The End of the World: The Ethics and Science of Human Extinction* ([7]), Leslie draws the conclusion that we should attempt to counteract the just-described probability shift by making the prior probability of Doom Soon very small. We should take measures to make the Earth a safer place, given that we live in 2001.

To discuss the Doomsday Argument in a precise way we need a minimum of probability theory. The proponent of the Argument asks us to consider two hypotheses, H_1 and H_2, Doom Soon and Doom Late, respectively. These are assigned prior probabilities p_1 and p_2, with $p_1 + p_2 = 1$. When new evidence E is taken into account, the probabilities shift in accordance with Bayes' rule:

$$P(H_i/E) = \frac{p_i.P(E/H_i)}{p_1.P(E/H_1) + p_2.P(E/H_2)}, \quad i = 1, 2. \tag{1}$$

Here $P(E/H_i)$ is the probability that E will occur if hypothesis H_i is true. $P(H_i/E)$ is the *a posteriori* probability that we should assign to H_i once we know that E actually occurred.

To see the seriousness of the doomsayer's conclusions, suppose that on the basis of our ordinary evidence we think that Doom Late has a probability of 99%, so $p_1 = 1/100$ and $p_2 = 99/100$. Let us denote by N_1 the number of human beings ever alive according to H_1; and let the corresponding number according to H_2 be denoted by N_2. The proponent of the Argument wants us to consider the case in which N_2 is very much larger than N_1, $N_2 = 10^6 N_1$, say. (The Argument's conclusions become the more pessimistic about the future of mankind the more optimistic the scenario H_2 is.) Finally, denote by n the number of people alive in 2001. (We have to assume that both hypotheses, H_1 and H_2, predict the same population figures at least up to and including 2001—otherwise we could falsify at least one of the hypotheses directly by inspection of past or present population numbers.)

The doomsayer now reasons as follows. I am an arbitrary member of mankind, and can consider myself as a random selection from all people ever alive. Therefore, according to H_i the probability of actually finding myself in 2001 is n/N_i. Of course, I was aware of living in 2001 all the time. Nevertheless, I can treat the fact that I live in 2001 as new evidence

E, because I did not take my place in history explicitly into account in my judgement of how probable H_1 and H_2 are. So I can use Bayes' formula to update my probability estimates, with $P(E/H_i) = n/N_i$. For the posterior probability of Doom Soon, once I have taken into account that I live in 2001, I thus find

$$P(H_1/E) = \frac{(1/100).(n/N_1)}{(1/100).(n/N_1) + (99/100).\{n/(10^6.N_1)\}} = \frac{1/100}{1/100 + 99/10^8} = 0.9999 .$$

Therefore, it becomes virtually certain to me that the extinction of mankind is imminent, after I have reflected on my existence now, in 2001.

2 Preliminary remarks about the Argument

In ([8, 9]) it was pointed out that the doomsayer should be careful to formulate his hypotheses H_1 and H_2 in a non-indexical way: if H_1 says 'Doom will strike soon', that would presuppose my existence (or the existence of the person who discusses the hypothesis) now, because 'soon' means 'a short time from now'. But if the hypotheses already contain the information E (my existence now, in 2001), subsequent conditionalisation on that information will not lead to a shift in probability (because $P(E/H_i) = 1$ in that case). So the Argument can only be taken seriously if the hypotheses have a form like: 'Doom will strike soon/late *after 2001*', without reference to my existence in 2001. More generally, the Argument as set forth above needs the supposition that I can be considered a random member of the collection of all human beings ever alive; that I can turn out to be any human in the history of mankind, and that all possibilities are equiprobable. That supposition provided the doomsayer with the justification for taking n/N_i as the probability, on hypothesis H_i, that I live in 2001. So the hypotheses H_1 and H_2 should contain no reference to the fact that I myself am one of the human beings living in 2001; they should leave it completely open when in human history I am actually living.

Even if the latter condition is fulfilled, it might be doubted whether it makes sense to consider myself a random member of the human race. After all, isn't it true that I couldn't live in any other period than I actually live in, given that I am a child of my parents, who are children of their parents, and so on? It could very well be that my existence here and now is the outcome of a deterministic process. Anyway, it seems obvious that there is no random mechanism in nature that can be held responsible for my turning up now in human history rather than at some arbitrary other instant of time. This is not a convincing objection against treating my birth date as a random variable, however. What is at stake in the Doomsday Argument is the probability that measures the strength of my belief—a 'subjective' probability. There is no contradiction between believing that some process is objectively deterministic, and assigning a subjective probability to its outcome that is neither 0 nor 1. Indeed, the date of the extinction of mankind itself may be the outcome of a deterministic process, and still it makes sense to consider our two doom hypotheses and their probabilities, because I do not know all relevant causal factors. Generally speaking, the most important applications of probability theory are situations in which we believe that underlying factors determine what will happen, but lack detailed knowledge about them. A situation in which I do not know my place in human history is certainly conceivable (see below for an example), in spite of the fact that my genetical background, etc., may determine exactly when in human history I live. There is nothing against quantifying my uncertainty in such a situation by means of a probability distribution. My actual situation is of course very different, because I am well aware that I live in 2001. This is reflected in different probabilities I should assign in my actual situation compared to a situation in which I lack information about my birth date (see below and section 3 for an extensive discussion of this point).

As explained, the core of the Doomsday Argument consists in an application of Bayes' rule. So it is clear from the outset that the Argument concerns the probability *shift* that occurs when new evidence is taken into account, rather than the absolute values of the probabilities. Consequently, the force of the Argument depends on the value of the prior probabilities that are to be used in Bayes' formula, Eqn. 1. If there were a reason for using a value for the prior probability of Doom Soon that is much lower than the probability we assigned before this whole issue of the Doomsday Argument arose, it could be that after the Bayesian probability shift the probability of Doom Soon is still not higher than we normally think it is. So it is important in the analysis of the Argument to critically consider the meaning and appropriate numerical value of the prior probability.

What the proponent of the Argument assumes, in order to arrive at his alarming conclusion, is that if we use our ordinary probability values as priors in Bayes' formula, the results are posterior probabilities that apply to our actual situation. This assumption is apparently based on the idea that in our actual probability evaluations we have not reflected explicitly on the fact that *we ourselves* live in 2001. Of course, we have used all kinds of information about the situation in 2001 (like information about the number of people alive in 2001, the state of technological knowledge, the risks of impending wars, etc.), but that *we*, instead of other human beings, live in 2001 has played no role in our calculations or intuitive probability judgements. So the prior probability to be used in Eqn. 1, which should pertain to a situation in which our place in human history is left open, seems to be the same probability we actually use without having thought about the Doomsday Argument ([3, p. 128]).

But is it really reasonable to follow the doomsayer here, and to use our actual probabilities as priors in Bayes' formula? Think about the knowledge situation this prior should correspond to. It is a situation in which we are asked to consider the hypothesis "the human race will become extinct relatively soon after the year 2001", and assign a probability to it. We are allowed to use evidence available to humans living in 2001 (the dangers of technological developments, warfare, etc.), while we must also assume it is consistent that we could subsequently get the information that we are cave-men living very, very long before 2001; or humans drastically changed by evolution, living millions of years later. This is a bizarre situation, very different from our actual one. It is not immediately obvious that the values of our prior probabilities in such a bizarre situation should be equal to our actual probabilities. In fact, there are arguments in the literature ([8, 10]) saying that it would not be irrational in this strange situation to choose our prior probabilities in such a way that subsequent conditionalisation on the information that we live in 2001 would result in our usual probabilities (the ones we actually assign in 2001, without thinking of the Doomsday Argument) as posteriors. However, the arguments adduced use ad hoc principles and can very well be contested (see [3, pp. 131-134]).

It is the purpose of the present paper to provide a more fundamental analysis of the Doomsday Argument. First, I will illustrate in what way, quite independently of the Argument, knowledge of one's place in history is relevant for the probabilities one assigns to the doom scenarios. In particular, if we attribute a probability p to one of the doom scenarios in our actual situation, in which we are aware of the year we are living in, consistency requires that we attribute a very different probability to the same hypothesis if we do not know what year it is, or decide to leave that information out of consideration. In fact, the new probability value should be exactly such that subsequent Bayesian conditionalisation on the information E about our birth date leads us back to the original value p. Of course, this is exactly what has to be expected. If we first discard a piece of information, and then put it back again, we should come back to the initial situation.[2] But I will attempt to make it clear exactly *how* knowledge about our place in history is relevant for the probabilities we assign to the doom hypotheses.

The observation about the triviality of the result of first removing information and adding it again later makes it clear that the Doomsday Argument has no force at all if we start from the probabilities that we assign in our actual situation, knowing our birth date, correct them so as to apply to the situation in which our place in history is open, and subsequently use the Argument with these corrected probabilities as priors. That would be moving in a circle. It also becomes clear that we cannot, on pain of inconsistency, use the same probability values both if we know and if we don't know what year it is. So we cannot use our actual probability assignments as priors in the Doomsday Argument, unless it is shown that these assignments are false, in the sense that they do not really apply to our actual situation but rather to the situation in which our place in history is completely left open. So what the proponent of the Doomsday Argument has to do before he can apply his Bayesian reasoning with any force is to convince us that the probabilities we actually use are not the right ones, but pertain to a different state of knowledge than our actual one.

Consequently, in order to judge the force of the Doomsday Argument only one question is important: can it be justified, independently of the Argument's Bayesian reasoning, that our actual probability assignments do not pertain to our actual situation, but rather to a different situation in which no knowledge about our place in history is available? Of course, if

[2]This remark also makes it clear that the asumption that our place in history is completely random, in the sense that a *uniform* probability distribution applies to our birth rank in human civilisation, is inessential. We could run the Doomsday Argument as well with an arbitrary non-uniform probability distribution, reflecting a *partial* loss of information about our place in history.

the answer to this question is affirmative, we should revise our probability judgements. This would apply not only to our probabilities for the doom scenarios, but to all probabilities we assign to possible future events.

So the second point of the paper concerns the justification of our actual probability judgements. This is an important issue, relating to the meaning of probability statements and their relation to empirical data; whereas the Doomsday Argument turns out to be essentially nothing but an exercise in applying Bayes' formula.

3 The relevance of time: an illustration

As we have seen, the prior probabilities to be used in the Doomsday Argument's Bayesian updating should pertain to a situation in which we do not know our position in time or decide not to use our information about when in human history we are living. These two possibilities boil down to the same thing from a mathematical point of view, because whether we actually do not possess information or consciously decide not to use it makes no difference in the formulas; the only thing that counts is whether the information is used or not. Let me illustrate the effect of lack of information about what year it is by a little story.

Suppose that I have been hypnotized and do not know when in human history I am living. I firmly believe that I may turn out to be any one of all individuals ever alive, be it Napoleon, a prehistoric troglodyte, or someone living in 130000 AD (if human civilisation will last that long); I think that every possibility is equally likely. Suppose further that I have been told that physical theory predicts a probability (objective chance) of 1/2 that the world will end in early 2002, given the state of the world in 2001; and an equally large probability that civilisation will go on for a very long time after 2001. It could be, for example, that a quantum mechanical spin measurement will be performed in the last minutes of 2001, with possible outcomes $+1$ and -1. If the outcome $+1$ is realized, a destructive device will automatically be activated that will annihilate the world within one hour; if -1 is the result, physical conditions then obtaining will make it certain that the human race will prosper for many thousands years. Our best physical theory, quantum mechanics, says that for each of the two possibilities the chance of being realized is 1/2, given the details of the experimental set-up. I now want to determine prior probabilities for my two hypotheses; probabilities that I can update when information about who I am will come in later. What probability should I assign to the possibility that the world ends in 2002 (Doom Soon)?

In order to relate the information about the physical chances to probabilities that quantify my degree of belief, I decide to follow the usual commonsense strategy, formalized by David Lewis as the 'Principal Principle' ([11]): as long as I have no information about the actual outcome of the chance experiment, and only know the value 1/2 of the chances, my 'reasonable subjective belief' in the occurrence of the outcome +1 should be given by the probability 1/2. More precisely, as long as all my evidence is compatible with the proposition about the physical chance (if the evidence is 'admissible'), the value of my 'credence function' will be given by the value of the physical chance.[3] The evidence is compatible as long as it does not imply anything about the actual outcome that is not already contained in the chance proposition. Because I know nothing about the actual outcome, it seems natural to take 1/2 as my probability for Doom Soon. When I subsequently learn that it is 2001 AD, and I update my probability using that information with the help of Eqn. 1, doom in 2002 becomes a virtual certainty to me, as explained in section 1.

So I seem to have justified the doomsayer's conclusion. However, the intuitive unreasonableness of that conclusion makes me pause for a moment. I reconsider my earlier argument, and it occurs to me that I have made a mistake in taking 1/2 as the value of the prior probability of Doom Soon. I have been assuming that it may now be any year in the history of mankind. But that implies that it may be a year far after 2002, the date of Doom Soon—namely if the Doom Late scenario represents the actual history. However, if the actual date is some year *after* the date of Doom Soon, this would imply that the outcome of the chance experiment was −1. If I found myself living in a year after 2002, the probability I would assign to the outcome +1 would therefore certainly not be 1/2; it would be 0. In Lewis' terminology, what I

[3] At this point the questions of whether or not the Principal Principle can be criticized, and of what the exact form of the admissability criterion should be, are not important. The purpose of this illustrative story is merely to show how even according to the usual way of handling probabilities, quite independently of the Doomsday Argument, information about the year I am living in is relevant. We shall return to the justification of equating physical probabilities and subjective probabilities in section 5.

have assumed about my position in time entails that the Principal Principle may not lead to the value 1/2 in the actual year I am living in, because the information about that year may be inadmissible; it may make it possible to infer the outcome of the chance experiment with certainty, so that the value of the objective chance becomes irrelevant. Obviously, I should be more careful than I was before in my assignment of probabilities to the two hypotheses. I should attempt to make a systematic calculation, taking the just-mentioned complication into account.

Let me denote the probabilities whose values I am going to determine by p_1 and p_2, pertaining to doom in 2002 and doom later, respectively. There are two possibilities for me to consider. First, I may be living before the date of Doom Soon, in which case the experiment is still going to take place and I can use my information about the physical chances for predictive purposes. In this case I will judge Doom Soon and Doom Late as equiprobable. Second, I may be living after the date of Doom Soon, in which case the physical chance is irrelevant for my belief in the truth of the hypotheses. In the latter case the probability that I will assign to hypothesis H_1 will be zero (because now I am absolutely certain that Doom did not strike in 2002). Consistency requires that I respect the rules of the probability calculus in calculating the probability that I have to assign to H_1 in my actual situation, in which I do not know when in human history I am living. It follows that my probability p_1 is equal to

$$\begin{aligned} &P(H_1|\,\text{I live before Doom Soon}).\,P(\text{I live before Doom Soon}) \\ &\quad +P(H_1|\,\text{I live after Doom Soon}).\,P(\text{I live after Doom Soon}) \end{aligned}$$

In other words:

$$\begin{aligned} p_1 = {} & P(\text{I live before the date of Doom Soon}).\,1/2 \\ & +P(\text{I live after the date of Doom Soon}).\,0 \end{aligned} \tag{2}$$

The probability that I should assign to the possibility that I live before Doom Soon (2002) is itself the sum of two probabilities, pertaining to two mutually exclusive cases: either H_1 is true, in which case I will certainly live before Doom Soon, or H_2 is true, in which case the probability that I, a human randomly selected from my long-lasting civilisation, live before the date of Doom Soon is N_1/N_2. (The two hypotheses, H_1 and H_2, must assign the same population numbers to all years before the date of Doom Soon to exclude the possibility of deciding between them by mere inspection of these figures.)

Things are simple now. It follows from what has just been said that the belief I attach to the possibility of finding myself in some year before the date of Doom Soon is represented by $p_1 + p_2.N_1/N_2$. Inserting this in Eq. (2), I obtain an expression that fixes the quotient of p_1 and p_2:

$$p_1 = (p_1 + p_2.N_1/N_2).1/2,$$

or, equivalently, $p_1 = p_2.N_1/N_2$. Therefore, the assumption that I lack all information about my place in history, together with consistency, compel me to adopt a value of the *a priori* probability of Doom Soon that is very much smaller than the *a priori* probability of Doom Late!

Inserting the result in Bayes' formula, I find for the *a posteriori* probability of Doom Soon, after I have taken the new information that I live in 2001 into account,

$$P(H_1/E) = \frac{p_1.n/N_1}{p_1.n/N_1 + (N_2/N_1).p_1.(n/N_2)} = 1/2.$$

So when I finally have the same information as I would have had without the intervention of the mesmerist, I think that H_1 and H_2 are equiprobable—just as I would have thought if I had not been hypnotized in the first place.

4 The general case

The result of the previous section shows that it is inconsistent to assign the same probabilities to the doom hypotheses both in the situation in which I do not know my place in history and in my actual situation, in which I know it is 2001. Probability theory specifies a definite relation between the probabilities in the two situations. This relation is such that the probabilities in my actual situation can be recovered from the probabilities in the situation in which I do not know the year I am living in by Bayesian updating with the evidence about which year it is. It is obvious that this point is general, and does not depend on the basis of my assignment of probability values in my actual situation—whether I use objective scientific results, astrology, or intuition does not matter because what is at stake is the *consistency* of my probabilistic reasoning rather than the concrete values of my probabilities. The equiprobability of the two hypotheses that was assumed in the example is of course also irrelevant.

Let me for the sake of completeness briefly rehearse the point for the general case. So consider the situation in which I decide, in full awareness that it is now 2001, to assign probabilities q_1 to H_1 (doom in 2002, Doom Soon) and q_2 to H_2 (Doom Late), respectively. It does not matter in what way the strengths of my beliefs, that are measured by these probabilities, have arisen. The only point of importance is that if my memory is suddenly partly erased, so that I no longer know the date, or if I decide to discard my information about the date, I have to revise my probabilities. In the new situation q_1 and q_2 are still the probabilities I *would* assign to the possible fates of the world if it were 2001 now; but in the new situation I do not know what year it is. Since I know that there are two possible scenarios for what can happen after 2001, I can deduce that the world will last at least until the beginning of 2002. However, q_1 and q_2 have predictive value only if I live *before* the date of Doom Soon. But I might be living *after* the date of Doom Soon (2002), in which case the probability I have to assign to Doom Soon is not q_1 but 0.

Making a calculation in the same way as before, I find that the probability I should assign to the Doom Soon hypothesis in the new situation, p_1, should satisfy

$$\begin{aligned} p_1 = {} & P(\text{I live before the date of Doom Soon}).\, q_1 \\ & + P(\text{I live after the date of Doom Soon}).\, 0 \end{aligned} \tag{3}$$

Furthermore, $P(\text{I live before the date of Doom Soon}) = p_1 + p_2.N_1/N_2$. Inserting this in Eqn. 3 I find that my new probabilities for the two hypotheses should be:

$$p_i = \frac{q_i.N_i}{q_1.N_1 + q_2.N_2}\,,\ i = 1,\, 2 \tag{4}$$

Subsequent conditionalisation on the evidence about which year it is now, by means of Bayes' rule with p_1 and p_2 as priors, leads to q_1 and q_2 as the posterior probabilities. These are the same probabilities I had assigned directly, knowing it is 2001.

The relevance of the knowledge about which year we are living in has become very clear now. If we have formed a belief about the probability of

Doom happening soon after 2001, given the situation of the world in 2001, we can use this probability value for a prognosis of Doom Soon *if we know that we live in 2001 or some earlier year*. If we lack information about the date, and may be living after the date of Doom Soon, we have to revise our probability estimates in the way explained above. This shows in what way we actually use our information about our own place in history, despite the doomsayer's contention that we do not.

5 The justification of our actual probability estimates

The analysis of the previous sections demonstrates that it is inconsistent to use the same probability values both in our actual situation and in the situation in which we lack information about when we are living. If we start with our actual probability judgements, subsequently correct them to obtain the probabilities for the situation in which we lack knowledge about the year we are living in, and finally apply Bayes' rule to add the missing information, we trivially end up with the probability we started with. But if we instead think that our usual probabilities really pertain to the situation in which we lack information about the date, Bayesian upgrading leads to a probability of Doom Soon that is much larger than our usual one. So the force of the Doomsday Argument depends entirely on the question of whether we are justified in thinking that our usual probability judgements pertain to our actual situation. The doomsayer's contention is that they do not; if the doomsayer is right it is a matter of course that we should revise our probabilities.

We already saw an argument for the doomsayer's claim, namely that in arriving at our probability assignments we have made no use of our knowledge that we live in 2001. Indeed, whether we rely on scientific results, astrological predictions or on other sources, our probability attributions are based on data available in 2001, but *not* on the thought that we ourselves happen to live in 2001 but in principle could have been living at a different epoch. Therefore, the doomsayer contends, what we really have done is justifying our probability assignments for the situation in which we do not know our place in history, instead of justifying them for our actual situation. However, this argument is not cogent. As we have seen in the previous sections, we *have* used the information that we live in 2001. If we had taken seriously the possibility that we could be living anytime in human history, we should have taken into account that we could be living *after* the date of Doom Soon. If we live after the date of Doom Soon our actual scientific, astrological, etc., probability assignments have no relevance; the probability of Doom Soon is then 0 instead. This should be taken into account in the calculation of new probability values, in the way explained before. The fact that we do not do so shows that we actually *do* use our information about the year we live in.

Still, the doomsayer might want to insist that we provide a solid foundation for our usual probability assignment. After all, he himself thinks it is rational to believe in a much higher probability of Doom Soon than the usual value. The first thing we have to observe in this connection is that as long as purely subjective probabilities are at stake, there can be no question of disproving the doomsayer's probability values. The doomsayer has the right to fancy any subjective belief he likes. If he wants to be consistent, he should not think that the same probability can be attributed to Doom Soon whether or not we have information about our place in history. But if he feels that the imminence of Doom is virtually certain, while he is aware of living in 2001, he should certainly quantify his degree of belief with a probability that is close to 1. (In that case consistency requires that he should assign a much smaller probability to Doom Soon in the situation in which there is no information about the year it is.) The dispute about the justification of our actual probability values therefore only becomes interesting if the probabilities are not purely subjective.

Lewis' 'Principal Principle', mentioned in section 3, has the aim of providing an objective basis for our probability judgements, so that they reflect 'reasonable belief'. The idea is to gauge our subjective probabilities by means of objective physical chances, provided by physical theory. Care should of course be taken not to overlook information that provides more detail about the events than contained in the statement about their physical probabilities—our credence function can only be set equal to the physical probability if the evidence is 'admissible'. For instance, if we have actually seen the outcome of a physics experiment, the physical probability of that outcome (given the earlier state of the physical system) is of no relevance for our believe in whether or not the outcome has occurred.

The typical basis of our actual objective, scientific, probability judgements is constituted by information about the state of the world now and about the previous history of the world, and by scientific theorizing using these data as input, while we know all the time that we live in 2001 and accept that we have no direct knowledge about the future. The example described in section 3 illustrates this general situation in a simple but adequate way. There it was supposed that our best fundamental physical theory supplied a probability value of 1/2 of Doom Soon (in 2002), given the state of the world in 2001. If we know that we are living in or before 2001, The Principal Principle can be applied and tells us to set our credence function equal to 1/2; this is the strength of reasonable belief advised not only by Lewis but also by common sense. The doomsayer has to provide an argument for his contention that this value is not the right one. As we have seen, merely reasoning that we have not used our information about our position in time is incorrect.

The doomsayer will be hard-pressed to supply such an argument. Scientific probability statements can be put to the test: if physics predicts that a physical system in state A will after a certain time show behaviour B with a probability of 50%, repetitions of the experiment, starting with systems in state A, should result in about half of the systems eventually showing behaviour B. This is an empirical claim, and predictions of this kind are routinely checked and confirmed in the laboratory every day. If we accept a physical theory that leads to probability statements, and want to assign probabilities in an objective way, it would be foolish not to use the physical probabilities as the values for our credence function. Indeed, accepting the physical theory involves the belief that the predicted probabilities will materialize as relative frequencies in the long run. If we decided to use different probability values, and based bets on them, we could be made to lose money consistently if the physical theory is right. Of course, this is exactly why the Principal Principle was put forward in the first place.

The example of section 3 is a case in point. The predictions of quantum mechanics concerning particles with spin have proved to be accurate in countless experiments. The theory predicts that if the spin, in a certain direction, of a spin-1/2 particle is measured the outcome will be either $+1$ or -1; both outcomes have a probability of 1/2. Of course, no experiments have ever been performed with a destructive device as described in section 3. But innumerable experiments have been done in which experimenters who were well aware of their position in human history measured and recorded the spin in a certain direction of a spin-1/2 particle. The quantum mechanical probabilities have always been verified as accurate predictions of the experimental frequencies (with the usual statistical fluctuations, of course). If I bet on other relative frequencies than 1/2, I will lose money. In fact, quantum mechanics also says that it does not matter whether an experimenter, conscious of the time or not, is present: the recording of spin values can be done mechanically, and the 50%-50% distribution will remain the same. So the *only* relevance awareness of my place in history can have is that if I know that I live *before* the time the experiment is performed, the quantum mechanical prediction is all I can use for my personal probability assignment to the possible outcomes of the experiment. If the experiment was performed in the past, it will in general be possible to find out what the outcome actually was, and the quantum mechanical probability will in that case become irrelevant for my personal probability, as discussed in sections 3 and 4. This is only different from the Doomsday situation described in section 3 in that in the latter case the mere fact that I exist after the date of Doom Soon is sufficient to infer the result of the experiment. There is no reason to assume that the quantum mechanical predictions, which have proved to be valid in an enormously wide range of circumstances, would suddenly lose their predictive value if an all-devastating bomb were made part of the chain of instruments coupled to the experiment.

The upshot of all this is that there are good reasons for relying on the usual objective, scientific, probability values. By contrast, all the doomsayer has to offer to convince us of the error of our ways is one incorrect argument (namely that in our actual probability assignments we have not used information about our place in history) and one irrelevant argument (the Doomsday reasoning itself, which is only about the *relation* between prior and posterior probabilities, and does not touch the question of what the values of our priors should be).

6 Conclusion

The Doomsday Argument does not support the conclusion its proponents think it supports; it does not make Doom Soon more probable than we ordinarily think. As an application of Bayesian reasoning the Argument is impeccable. But Bayes' formula by itself does not imply anything about the values of prior or posterior probabilities; it is only about the relation between the two. To make it plausible that we should revise our usual probability assignments, the doomsayer should convince us that these usual probability assignments are false; and he should do so independently of his Bayesian reasoning. I have argued that there is nothing in the doomsayer's argumentation that comes near to this.

References

[1] J. Leslie, *Universes*, Routledge, London, 1989.

[2] J. Leslie, Doom and Probabilities, Mind 102 (1993) 489-491.

[3] N. Bostrom, Observational Selection Effects and Probability, Doctoral Dissertation, London School of Economics, 2000. Available at http://www.anthropic- principle.com/phd.

[4] N. Bostrom, The Doomsday Argument is Alive and Kicking, Mind 108 (1999) 539-550.

[5] N. Bostrom, Investigations into the Doomsday Argument, preprint available at http://www.anthropic-principle.com/preprints/inv/investigations.html.

[6] N. Bostrom, The Doomsday Argument: a Literature Review, preprint available at http://www.anthropic-principles.com/preprints/inv/investigations.html.

[7] J. Leslie, *The End of the World: The Ethics and Science of Human Extinction*, Routledge, London, 1996.

[8] D. Dieks, Doomsday—Or: The Dangers of Statistics, The Philosophical Quarterly 42 (1992) 78-84.

[9] D. Dieks, More About Doom, The Mathematical Intelligencer 15 (1993) 4-5.

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

[11] D. Lewis, A Subjectivist's Guide to Objective Chance, in: D. Lewis, *Philosophical Papers, Volume II*, Oxford University Press, Oxford, 1986, 83-132.