Originally published in *Mind*, Vol. 102, No. 407
July 1993

# *Probability Theory and the Doomsday Argument*

**WILLIAM ECKHARDT**

## *I*

John Leslie has developed a statistical argument that suggests the end of the human race may be much closer than we generally suppose, even in our current mode of ecological pessimism. (Leslie 1990a, 1990b, and 1992.) This argument, which he attributes to Brandon Carter, can be summarized as follows.

Among all people who have ever lived, our current birth-order rank is something like 50 billion. If humanity is to continue as long as we usually suppose, then we shall have a quite low rank among all who ever live. According to Leslie, this should be considered unlikely; it is more probable that we have an average rank among all who ever live, in which case doomsday will be much sooner than we generally suppose. In effect, one's birth is treated as though it were a random drawing from a lottery of unknown size, consisting of all humans who ever live. One could use one's own rank in the successive drawings from such a lottery to estimate the size of the entire lottery pool. This suggests that the entire human lottery is not large relative to our rank, i.e., doomsday is likely to be relatively soon.

With respect to a true lottery, this reasoning is an uncontroversial application of Bayes' theorem. But the question remains whether such reasoning applies to the likely fate of the human race.

## *II*

Leslie reviews a host of possible objections to the Doomsday argument. However, I believe he fails to locate the crucial issue upon which the whole edifice rests, namely that we should consider ourselves as *random* members of the total human population. For instance, in countering the objection that the argument's reference class, all human beings, is wrongly chosen, and should instead be all conscious observers (Leslie 1992, p. 530), he points out that only human survival is at issue, therefore, only humans need be taken into consideration. However, do we have better reason to believe that we are random humans than random vertebrates or random social animals? Can the same item be random in all these classes? Isn't a *random* human a rather exceptional vertebrate? Isn't a *random* social animal much more likely to be an insect than a human? These questions press, irrespective of their relevance to the issue of human survival.[1]

My primary purpose is to examine the statistical underpinnings of the Doomsday argument, but preliminary to this I would like to disentangle the problem from certain perennially unresolved philosophical issues with which it has come to be associated. As long as the validity of the Doomsday argument is made to hinge on whether the future is open or fixed, or whether the future is fully implicit in the present, we can rest assured we are not going to settle the question of the argument's validity.

Leslie asserts that both the truth of indeterminism and its importance modulate the impact of the Doomsday argument. He claims that an open future "reduces the power of Carter's reasoning, instead of destroying it" (Leslie 1992, p. 537), but gives as explanation only the possibility that indeterminism may not matter much to human survival. Yet all the key ingredients of the argument – Bayes' theorem, our birth rank, and our prior expectation of doomsday – are such that one cannot say why determinism should make a difference to them. Leslie does not mention determinism either in his central presentation of the argument or in numerous collateral examples. It is unclear what step of the argument a failure of determinism is supposed to weaken.

In fact, the issue of determinism is a red herring. Consider an urn lottery in which balls are admitted to the pool according to the dictates of a chance device. The last ball to be admitted and hence the pool size is decided probabilistically. The chance device might consist of a systematic algorithm involving the decimal expansion of $\pi$, of tossing a series of coins in a Newtonian universe, or of consulting quantum events of appropriate probability. The first two procedures are strictly deterministic; the third, according to the prevailing opinion of theoretical physicists, is indeterministic; yet, one would make

[1] Vagaries of the reference class also cloud the issue of what constitutes empirical confirmation of the Doomsday argument. Suppose in one hundred years people stop reproducing in the way that is currently customary and for the next million years the human race consists of human brains inside of robots. Is this a confirmation of the doomsday reasoning because among all flesh and blood humans we then have average rank? Or is it a disconfirmation demonstrating the need for some reason as to why we chanced to be born so early in the history of the race that we are not brains inside of robots?

identical statistical inferences in each case, provided the probabilities generated were, in each case, the same.

Assuming determinism assures a rigid connection of any future event to the ensemble of *all* present events, but surely this does not imply a determinable connection of a given future event to a *specific* present event. It is quite consistent with determinism that the correlation of doomsday and the present event consisting of one's ranking in the human population to date be zero or vanishingly small.

The determinist believes the future to be implicit in the present, but cannot utilize this belief to make a prediction. (The determinism/indeterminism distinction is notoriously devoid of practical consequence.) Determinist and indeterminist are on exactly the same footing when it comes to making probabilistic inferences. The practicing statistician need not be concerned with questions of whether physical process is ultimately deterministic or whether the future is open or fixed. If the statistical reasoning underlying the Doomsday argument is unique in this regard, the matter requires further elucidation.[2]

## *III*

We proceed with an analysis of the Doomsday argument on the basis of probability theory.

*Rank* is birth-order rank since the beginning of mankind. Let $Dm(d)$ mean that $d$ is the rank of the last human who will ever live (doomsday). Denote by $Samp(r)$ that a specified one-shot sampling procedure results in the choice of a person of rank $r$.

The sample rank used in the argument is supposed to be our own. For definiteness, we can use $n$, the rank of a baby born now. What consequences does $Samp(n)$ have for the distribution of $Dm(d)$? According to Bayes' theorem:

[2] The impression that determinism is relevant to the Doomsday argument may be motivated by the following somewhat inchoate reasoning: if I am to be a random member of the total human population, my expected rank needs to be the average human rank, but the average human rank depends on how many come after me. If the population to come after me were subsequently increased, say through the intervention of a benevolent angel, it would be unreasonable to suppose that my expected rank would be retroactively increased; we can only conclude that such unforeseeable additions to the human pool would compromise my status as random. The assumption of determinism serves to keep this potentially unruly future under control.

(1) $$P(Dm(d) \mid Samp(n)) = \frac{P(Samp(n) \mid Dm(d))\pi(d)}{\sum_{r=0}^{\infty} P(Samp(n) \mid Dm(r))\pi(r)}$$

where π(*d*) is the prior distribution of doomsday, prior, that is, to the application of the Doomsday argument.

The probability $P(Samp(n) \mid Dm(d))$ depends on the sampling procedure used to generate *n*. Consider the following two examples.

In the first, sampling is equiprobable sampling from the collection of all people past, present, and future. Denoting this purely random sampling procedure by *Rand*(*n*), we have:[3]

(2) $$P(Rand(n) \mid Dm(d)) = \begin{cases} 0, \text{if}(n > d) \\ \frac{1}{d}, \text{if}(n \le d) \end{cases}$$

Then the posterior probability of *Dm*(*d*) given sample *n* is:

(3) $$P(Dm(d) \mid Rand(n)) = \frac{\frac{1}{d}\pi(d)}{\sum_{r=n}^{\infty} \frac{1}{r}\pi(r)}$$

And in this case, exactly analogous to a lottery, there is a shift to the earlier in the distribution of doomsday. Note that the distribution of *Rand*(*n*) depends on *d*.

However, the sampling arrangement in this example cannot truly be analogous to that of the Doomsday argument. In sampling *equiprobably* from a pool, only part of which currently exists, it is essential that one not invariably succeed in obtaining a sample item. Equiprobability entails that in some instances the sample ought to be one of

[3] Leslie's argument suggests in places that it is $P(r \le Rand(n) \mid r \le d)$ that is the relevant one. In the case of equiprobable sampling

$$P(r \le Rand(n) \mid r \le d) = n/d = nP(Rand(n) \mid Dm(d));$$

this extra factor *n* cancels out during normalization yielding the same result.

the nonexistent items, in which case the procedure ought to yield a null result.[4] A procedure that invariably yields an existent item cannot be equiprobable sampling, since in that case nonexistent members of the pool could not be receiving appropriate weight. Yet the sampling procedure employed in the Doomsday argument invariably yields a result – a human rank current at the time of the argument's discovery. Hence, this cannot be equiprobable sampling from an ensemble only part of which currently exists.

The second example is motivated as follows: in generating a sample from an ongoing process of *unknown* termination, one typically would not expect the sampling distribution to be affected by what happens to the process *after* sampling. The only effect the unforeseeable termination should have in this regard is to preclude any new sampling. This would be the case, for instance, in sampling from admissions to the urn lottery described above.

To reflect this feature let *Samp*(*r*) represent a sampling procedure that does not depend on the value of *d*, i.e., *Samp*(*r*) is the restriction to $r \le d$ of a distribution that is independent of *d*. Applying Bayes' theorem yields:

$$(4) \qquad P(Dm(d) \mid Samp(n)) = \begin{cases} \dfrac{P(Samp(n) \mid Dm(d))\pi(d)}{\sum_{r=n}^{\infty} P(Samp(n) \mid Dm(r))\pi(r)}, \text{ if } (d \ge n) \\[2ex] 0, \text{ if} (d < n) \end{cases}$$

For $d \ge n$ the numerator of this expression equals:

$$(5) \qquad \frac{P(Samp(n))\pi(d)}{\sum_{r=n}^{\infty} \pi(r)}$$

and the denominator equals *P(*Samp(n)). Therefore:

$$(6) \qquad P(Dm(d) \mid Samp(n)) = \frac{\pi(d)}{\sum_{r=n}^{\infty} \pi(r)}$$

[4] In sampling equiprobably from a population in which only some members are available as potential samples, the probability that the procedure fails to produce an item has to equal that fraction of the entire population that is unavailable. This is why truly random sampling from the entire human pool would be so informative. Its failure rate would be a reliable guide to the size of that portion of the population that is currently nonexistent.

The last expression is merely the restriction of the prior distribution of doomsday to cases greater than or equal to the sample; this reflects the fact that the sample does tell us one thing about doomsday, namely, that doomsday cannot come earlier than the sample itself. Thus, in the typical case wherein the sampling procedure does not depend on *d*, the information contained in *Samp*(*n*) yields absolutely no shift of the prior distribution of doomsday.

This second example shows that, for application of Bayes' theorem to result in any shift of our expectations of doomsday, there has to be some sort of probabilitistic dependence of the sample rank, *n*, on the doomsday rank, *d*. Is such dependence plausible?

## *IV*

It may be objected that the probabilistic independence of the sampling procedure from *d* is not perfect. Independence is generally an idealization, and this case is no exception. It is plausible that the quantitative probability tools necessary for the formulation of the Doomsday paradox are also needed for the formulation of, say, quantum mechanics, which itself impinges on the likelihood of an early doom. In other words, there may be correlation between *Samp*(*n*) and *Dm*(*d*) owing to some factor *f* whose presence hastens both the discovery of the Doomsday argument and doomsday itself. If so, we can assure the independence of the sampling distribution from *Dm*(*d*) by taking *f* into consideration.

Denote by *DmArg*(*n*) the procedure that draws sample rank *n* upon the discovery of the Doomsday argument. Factor *f* should "screen-off" the correlation of *DmArg*(*n*) with *Dm*(*d*), that is:

(7) $$P(\text{DmArg(n)} \mid \text{f and Dm(d)}) = P(\text{DmArg(n)} \mid \text{f}), \text{ for } n \geq d$$

Applying Bayes' theorem, we can conclude that:

(8) $$P(\text{Dm(d)} \mid \text{f and DmArg(n)}) = P(\text{Dm(d)} \mid \text{f})$$

and as in the last case, the information that *DmArg*(*n*) has no effect on the distribution of *Dm*(*d*) given *f*, other than to assure that $d \geq n$.

If correlation of *DmArg*(*n*) and *Dm*(*d*) also reflects direct causal influence, then the screening-off condition does not apply. (As an illustration, consider that the

Doomsday argument might prompt people to strive to lower the probability of doomsday.[5]) Such direct causal influence is likely to be exceedingly slight, and it is furthermore sufficiently distant from the intentions of the argument that it may be safely disregarded.

*V*

The foregoing considerations all support the contention that the sampling procedure used in the Doomsday argument is of the kind that engenders no shift in prior estimates of doomsday. There may exist a plethora of reasons for supposing the human race to be doomed, but our own birth rank in the total human population cannot reasonably be counted among them.

WILLIAM ECKHARDT

## REFERENCES

Leslie, John. 1990a: "Risking the World's End". *Interchange*, 21, 2, pp. 49-58.
_____ 1990b: "Is the End of the World Nigh?". *The Philosophical Quarterly*, 40, 158, pp. 65-72.
_____ 1992: "Time and the Anthropic Principle". *Mind*, 101, 403, pp. 521-540.

[5] Along these lines, I heartily endorse Leslie's concerns about accelerator experiments' tripping a descent from the false vacuum to some lower state (Leslie, 1990a, pp. 54-55; 1992, p. 525). Such possibilities deserve to be better known. Evidently there is something so clean about such complete annihilation, analogous to vanishing in a genie's spell, that people do not feel the emotional impact that they do from contemplating slower deaths by nuclear holocaust or ecological disaster.