Doomsday: Two Flaws

Mike Lampton

Space Sciences Lab, UC Berkeley

mlampton@berkeley.edu

Abstract

Here I argue that the much-discussed Doomsday Argument (DA) has two flaws. Its mathematical flaw stems from applying frequentist probability where statistical inference is needed. Its conceptual flaw is assuming that Copernican uniformity applies to a collection of ranked people and their ideas. I conclude that the DA has no predictive power whatsoever.

Keywords— Doomsday, statistical inference, Bayes theorem

1 Introduction

The "Doomsday Argument" (Carter 1983, Gott 1993, Leslie 1996) applies a kind of Copernican principle that today's humankind does not occupy any special place in our long-term biological time span. With this hypothesis – uniform random *now* – there is only one parameter to be estimated, namely the total number N of individuals who will ever exist. For any given size N, this argument leads to a statistical confidence region for our future prospects, based entirely on the cumulative rank R of individuals as of today: in only 1% of trial universes under this hypothesis would inhabitants find $R < 0.01N$; 99% of all such trial runs would have $R > 0.01N$. Today our rank $R \approx 10^{11}$, a reasonable finding if $N = 2 \times 10^{11}$ or even 10^{12}. So we should be 99% confident that $N < 10^{13}$, and doomsday is highly likely in the not-too-distant future.

A concise statement of the DA was posed by John Leslie (1996) p.1:

> "We ought to have some reluctance to believe that we are very exceptionally early, for instance in the earliest 0.001 percent, among all humans who have ever lived."

Its strengths and weaknesses have been discussed by many authors (Dieks 1992, Kopf et al 1994, Leslie 1996, Hanson 1998, Korb and Oliver 1998, Oliver and Korb 1998, Bostrom 1999, Monton and Rousch 2001, Olum 2002, Sober 2003, Leslie 2008, Leslie 2010, Northcott 2016, McCutcheon 2018); here I shall simply fix on its two major errors.

2 The Mathematical Flaw

To begin, where are we now? Our present rank R is about 10^{11} and is growing as I show in Figure 1 below.

I will now apply Bayesian inference by decomposing the universe of possibilities using two numbers. Let each such civilization terminate with a finite cumulative number of members N, and for each civilization let each member be assigned a birth rank R with $1 <= R <= N$. Next I will apply the clueless assumption: my own personal rank R is uniformly distributed within the range $1 <= R <= N$, and given N, my particular probability of having any one ticket number is $1/N$.

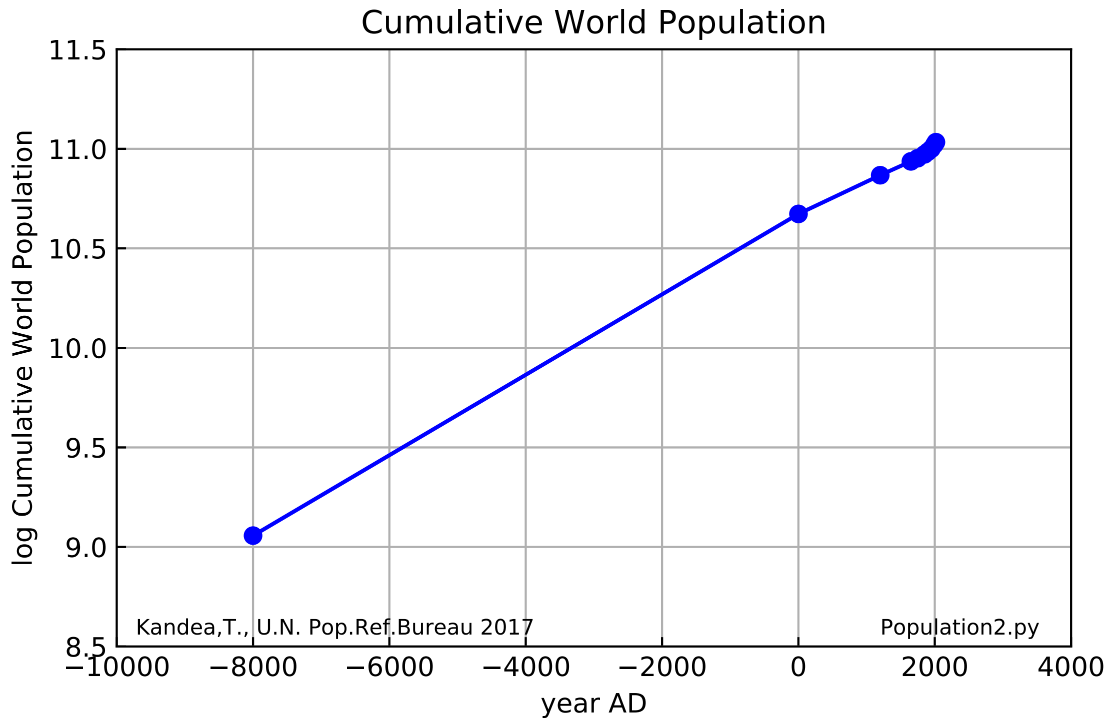

Figure 1: Cumulative number of persons versus time, worldwide

Finally I will adopt logarithmic scoring (McCutcheon 2019) and calculate the Bayes relative likelihood of two classes of futures: QuickDoom "QD" $10^{11} < N < 10^{12}$, versus DistantDoom "DD" with $10^{17} < N < 10^{18}$. These two classes of futures have, for each N and priors $p(X)$, a relative likelihood of

$$\frac{L(QD)}{L(DD)} = \frac{p(QD) \cdot \sum_{10^{11}}^{10^{12}} \frac{1}{N}}{p(DD) \cdot \sum_{10^{17}}^{10^{18}} \frac{1}{N}} = \frac{p(QD) \cdot [ln(10^{12}) - ln(10^{11})]}{p(DD) \cdot [(ln(10^{18}) - ln(10^{17})} = \frac{p(QD)}{p(DD)} \tag{1}$$

For equal size log bins, the statistics cancel out. The relative likelihood is just the ratio of our priors! I must be somewhere, but until I get more data, all I know about the size of my cohort is that it exceeds my rank. It could be small, it could be large.

Where does the DA fail, exactly? Originally it was posed as a purely frequentist probability issue (given N, it is astonishing to have drawn a rank $R < 10^{-6}N$) and that is now often recognized as being an invalid circular argument. The more recent DA formulations use Bayesian inference, as they should, but with flaws (Oliver and Korb 1998 examine many of these). The simplest of the Bayesian DA formulations is to compare one simple hypothesis, say QD=10^{12}, against another simple hypothesis, say DD=10^{18}. No sums needed:

$$\frac{L(QD)}{L(DD)} = \frac{p(QD) \cdot P(R|QD)}{p(DD) \cdot P(R|DD)} = \frac{p(QD) \cdot 1/QD}{p(DD) \cdot 1/DD} = 10^6 \cdot \frac{p(QD)}{p(DD)} \tag{2}$$

For very reasonable equal priors, this tells us that QD is a million times more likely than DD. Convincing, until you ask about the likelihood contributions from the 10^{12} nearby neighbors of QD and the 10^{18} nearby neighbors of DD. Shouldn't they count, as well?

One faulty way to include those neighbors is to average them over their respective zones. To this I would reply that averaging is inappropriate: likelihoods add since each added simple hypothesis contributes another way to obtain the observed rank.

Another faulty way to include those neighbors is to assume that the prior of each future N is known in advance from some previous measurement, and was found to equal to $p(N) = 1/N$. This prior can then be installed into the composite summations, eliminating the group priors $p(X)$ in equation 1. This gives us a likelihood ratio of

$$\frac{L(QD)}{L(DD)} = \frac{\sum_{10^{11}}^{10^{12}} \frac{1}{N^2}}{\sum_{10^{17}}^{10^{18}} \frac{1}{N^2}} = \frac{1/10^{11} \ - \ 1/10^{12}}{1/10^{17} \ - \ 1/10^{18}} = 10^6. \tag{3}$$

So the QD scenario is a million times more likely than the DD scenario, just what we put in with the priors.

My view is that this DA misunderstands priors. The Bayesian process is to acknowledge that previous information exists about our two chosen alternative hypotheses. From those priors we use conditional probabilities to update our opinion in view of new data. Since there is no prior information on what N might be, we are obliged to regard all $p(N)$ values as equal, and so we return to equation 1.

View from Above

Here's a way to visualize both the frequentist probability and the Bayesian inference at once. In Figure 2 below I show a finite portion of all possible setups and all possible outcomes.

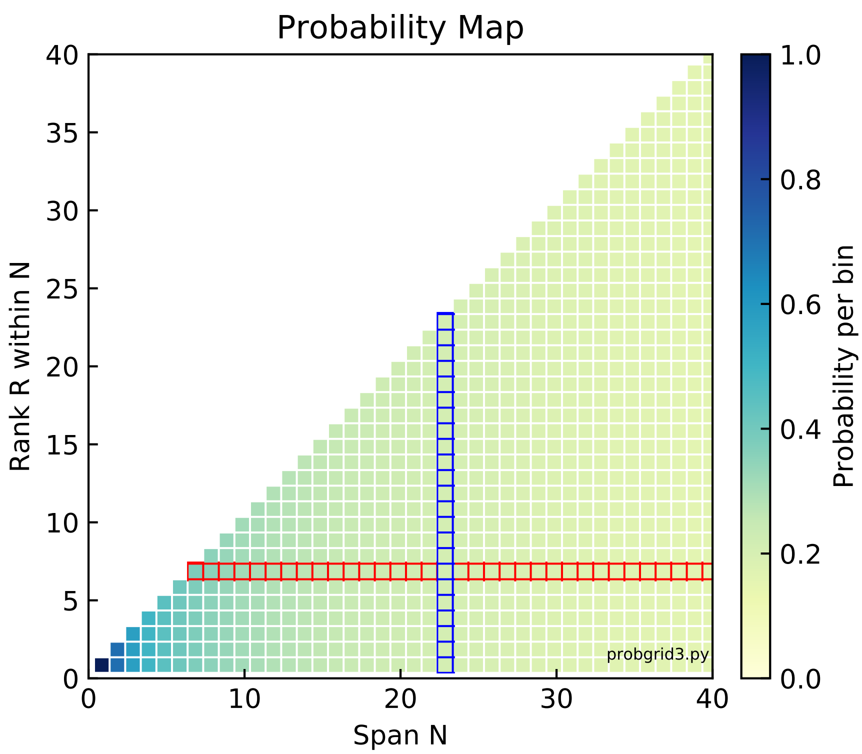

Figure 2: Map of conditional probability $P(R|N)$. For each population size N (any vertical bar) , there are N possible ranks and each rank has a probability of $1/N$. This gives a simple finite confidence region and leads to Doomsday. But for each rank R (any horizontal bar) there are an infinite number of population sizes that contribute to its likelihood. Every decade in N beyond R contributes a likelihood of 2.303. This equality defeats Doomsday.

The frequentist view (vertical blue band) is a single column in this diagram: given a cohort size N, the probability of each ranked individual is $1/N$, equal for all ranks. Being finite, it has simple confidence regions: you are 99% confident of not drawing a first percentile number. The Bayesian view (red horizontal band) is a single row in this diagram: we have drawn a 7, so what may we infer from this? *This is no longer a finite problem.* If all N values have equal *a priori* probabilities, we gather a likelihood of $1/7$ from $N = 7$, plus $1/8$ from $N = 8$, plus $1/9$ from $N = 9$, etc., all the way to infinity. From this harmmonic series we may now extract our composite hypotheses (for example, QD vs DD) for comparison. The ratio of any pair of decades or octaves above the observed ranks is 1. This is Rev. Bayes way of telling us that beyond our rank R, there is no information here about the size of our cohort from a single sample.

The UrnFiller vs the UrnGuesser

The UrnFiller (let's call her Jill) has two Urns (she calls them A and B) and into A she puts ten balls, numbered 1 through 10, and stirs them up. Into B she puts 100 balls, numbered 1 to 100, and stirs them up. She is good at probability, and calculates that the chances of some future guesser will draw a 7 is $P = 0.5 * 0.1 + 0.5 * 0.01 = 0.055$ exactly: two possible choices between urns and then $1/10$ or $1/100$. She is right. This is pure frequentist probability; no Bayesian inference needed.

The UrnGuesser (let's call him Kent) has a different problem. He knows that Jill never lies, and is told the urns contain 10 and 100 balls. The urns of course look exactly the same. He picks an urn. He is not allowed to weigh it. He picks a ball. It is 7. Now what? There was a guaranteed 7 in each urn! So he has learned nothing about his choice of urns. Might have been the heavy urn, or the light urn. (Had he picked a 61, he would have nailed it, but alas that didn't happen.)

But wait! The UrnGuesser has learned about Bayesian inference, so applies it like this, to compare his relative likelihoods of two narrow categories Urn10 versus Urn100:

$$\frac{L(10)}{L(100)} = \frac{prior(10) * P(7|10)}{prior(100) * P(7|100)} = \frac{prior(10) * 0.1}{prior(100) * 0.01} \tag{4}$$

With this, he argues that it is ten times more likely that his chosen urn was the 10 than the 100. Those two sevens are ten times more dilute among 100 than they are under 10. Being dilute means they are not going to be found so often. A universe of ball-pickers should agree. Between these two narrow categories, Bayesian inference gives us perfect agreement with the UrnFiller's frequentist conclusion. Problem Solved. QED. End of discussion.

But wait! What if *there is no UrnFiller*? and *what if there is only one urn*? The UrnGuesser picks a ball. It is a 7. How many balls were in the urn? This is a different problem, not a simple choice between two completely defined alternative narrow hypotheses, but between members of an infinite continuum of urns. It is this switch that forces us to abandon frequentist probability and adopt Bayesian likelihood methods.

The general schema is to identify two groups of distinct, non-overlapping hypotheses, add up the conditional probabilities of each simple hypothesis in each group, and compare their likelihoods. Like this:

$$\frac{L(A)}{L(B)} = \frac{\Sigma prior(A) * P(7|A)}{\Sigma prior(B) * P(7|B)} \tag{5}$$

To do this we must first decide what our groups will be. Having found a seven, it is certain that there were at least seven balls, but beyond that, the sky's the limit. Let's create two alternative composite hypotheses: H70 spanning $7 < N < 70$ versus H700 spanning $70 < N < 700$.

For each N, before drawing any balls, the conditional probability of getting a 7 is simply $1/N$. Since we have no reason to prefer one N value over another, the *a priori* factors are come outside the summation and cancel. , Our integrated conditionals are just $ln(10) = 2.302$ in both cases. So the likelihoods of the two cases are equal. We have no reason to prefer one decade over the other.

Next, after drawing one ball, all decade spans above its rank R remain equal, and the *a priori* factors are equal, so the likelihood ratio is 1.0 for all test decades beyond R. From one sample we learn only that we may rule out $N < R$ which we knew already. Again, we have no reason to prefer one decade over the other, provided that both decades lie beyond R. We still know nothing about the size of our cohort.

The Panzer Battle Commander vs the Allied Defense Commander

Our PBC knows he has fielded 1000 tanks and (unfortunately for him) each tank carries its serial number, 1...1000. If his tank #77 were captured in battle, that was in his first decile, but no surprise: it had to be in some decile, and all ten deciles are equally probable. But if there were a second capture, say tank #34, also in the first decile, that would be improbable: only a 10% that two captures would share a common bin. Extending this, the probability of N captures from the first decile is 10^{-N}, and the probability of N captures all from a common decile is 10^{1-N}.

Our ADC faces a very different problem. He does not know N. Had he captured a usefully large number K of tanks and found a maximum serial number M, he would reasonably suppose that his captures span almost the entire production run and figure that N is just a bit bigger than his M. In the large-K approximation, he would write $\langle N \rangle \simeq M + AverageGap = M + M/K$. This method fails for small K (see Wikipedia 2019): the variance in the $\langle N \rangle$ estimator is infinite for $K <= 3$; indeed the mean $\langle N \rangle$ is infinite if $K <= 2$. One tank tells him squat about the size of its cohort.

My Reply to John Leslie's "Objection f"

In his thoughtful compendium of views on the DA, Leslie (1996, page 23) warns us:

> "Don't object that there would be more chances of being born into a long-lasting human race, and that these would precisely compensate for being born early in the history of that race. The answer to this is that there would be nothing automatically improbable in being in a short-lasting human race. (Imagine that only ten people will ever have been born. Ought you to be specially surprised at finding yourself among the ten? No, for only those who are born can ever find themselves as anything.)"

My view is that this objection conflates a fact with a distribution:

Draw one sample. Its observed rank is R. Now R has graduated from being a distributed statistic to being a fact. We now have a firm lower limit on the population size $N >= R$, but still no upper limit. For each remaining unobserved sample of rank M we know $M \neq R$ and we can now write $P(M|N) = 1/(N-1)$: M remains uniformly distributed on N except for the point $M = R$. Yes, having drawn one sample, you still have a 50/50 chance of being in the big urn cohort vs. the small urn cohort since one sample tells you nothing about the size of your cohort. But the $1/N$ or $1/(N-1)$ densities continue to apply to the remaining draws as yet unseen.

My major point is that to adopt a single N, then obtain (say) a 99% confidence region for R, then from our observed R to regard that $100 \times R$ as a 99% confidence maximum on N, is a circular argument. It must be replaced with a confidence region obtained from Bayesian inference, and with just one sample that region has no upper limit.

3 The Uniformity Flaw

Figure 3 presents a linear quasi-Copernican sequence of all humanity (discussed for example by Cirkovic and Balbi 2019). Suppose there are ten trillion (10^{13}) individuals, ranked by birth. Uniformly throughout, people formulate ideas randomly (spots). Any idea can appear at any time, without any particular historical relationship to the many other ideas represented in the diagram. Some are good ideas, some are bad. One of them (red) might be the Doomsday Argument.

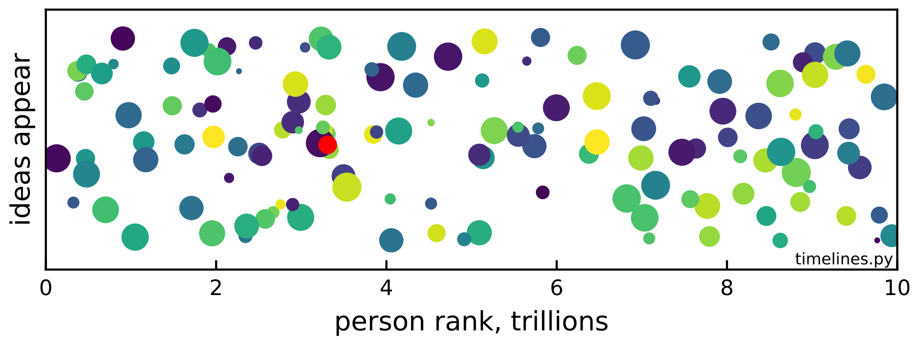

Figure 3: A cartoon illustrating a statistically uniform distribution of people and their ideas.

Figure 4 shows that same sequence on a logarithmic scale. But here I do not regard ideas as unrelated. They build on previous ideas; they become confirmed or modified or rejected. So, I plot knowledge as a curve of networked ideas mostly enjoying a steady rise[1] in breadth and confidence. In particular, the discovery of the DA depends on writing, mathematics, thought experiments, and probability. At our present rank we have recently entered the Zone: we have invented the DA, we are discussing it, and we may soon outgrow it.

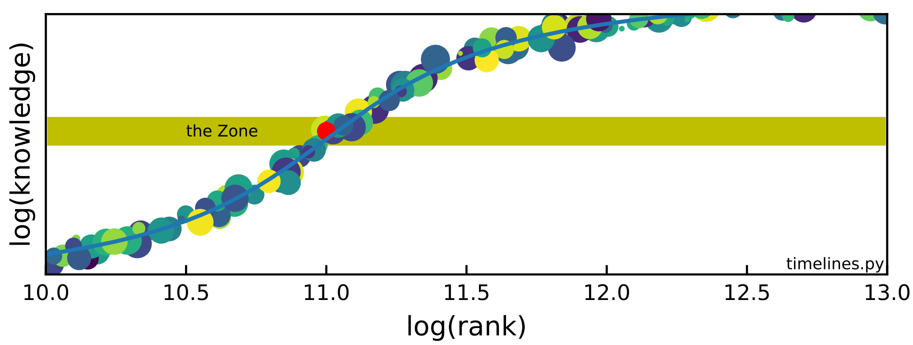

Figure 4: Logarithmic scale cartoon illustrating the growth of knowledge as an accumulation of previous ideas and their interrelationships.

A dedicated Doomsdayist would immediately object that our Zone is at rank $R \simeq 10^{11}$ because that is reasonably situated among $10^{11}..10^{12}$ civilizations, whereas long-lived Distant Doom civilizations, (say) those with $R \simeq 10^{17}$, would statistically be unlikely to find their Zones around 10^{11}: they are far more likely to be encountered at random locations throughout their much larger spans.

[1]This rise need not depend on any improvements in human intelligence. Gould and Eldredge (1977) describe evolutionary trends as punctuated equilibria: long periods without significant genomic change, interrupted by relatively rapid shifts in their properties. Quite likely, human intelligence has followed this *punk eek* pattern. Bostrom (2009) compares long-term static intelligence versus various slow and rapid improvements. One extreme is the singularity (Ulam 1958; Good 1965; Vinge 1993; Sandberg 2009) that propels human capabilities far beyond those envisioned today. My take on these possibilities is that — even without them — growth in aggregate human knowledge will assure a much improved likelihood of overcoming future crises.

To this, I have two responses. First, the Zone is systematic, not random, and its arrival rank is set by the historical development of essentials like writing, teaching, and mathematics, all of which come from the past. They are utterly independent of future developments. Statistical methods do not apply. Second, it is an error to apply conditional probability $P(R|N)$ to a question demanding inferential likelihood $L(N|R)$.

4 Conclusions

Two extreme cases – a modestly future-human and an outrageously future- or post-human scenarios — cannot be mathematically distinguished by knowing only our present population size. Moreover, the growth in human knowledge renders the assumption of constant disaster probability untenable. These thought experiments leave little doubt (in my mind, at least!) that the conceptual setup of enumerating a sequence of equal-probability bins ignores the demonstrable fact that advancing civilizations become more aware of existential difficulties, and thereby modify their future outcomes. Humanity's passenger train definitely starts out limited in size, but might outgrow any such limit as time progresses.

Acknowledgment: I gratefully acknowledge having illuminating discussions with John Leslie and Ken Olum on the many facets of the DA. My errors are of course mine alone.

References

Bostrom, N., "The Doomsday Argument is Alive and Kicking," Mind v.108, pp. 539-550 (1999).

Bostrom, N., "Future of Humanity," New Waves in Philosophy of Technology, eds. Jan-Kyrre Berg Olsen, Evan Selinger, and Soren Riis (New York: Palgrave McMillan, 2009).

Carter, B. "The anthropic principle and its implications for biological evolution," Phil. Trans. Royal Soc London A310, pp.347-363 (1983).

Cirkovic, M., M., and Balbi, A., "Copernicanism and the typicality in time," IJA (2019).

Dieks, D., "Doomsday — or: the Dangers of Statistics,", Philosophical Quarterly v.42, Issue 166, pp. 78-84 (1992).

Good, I.J., "Speculations Concerning the First Ultraintelligent Machine", Advances in Computers, vol. 6 (1965).

Gott, J.R., "Implications of the Copernican Principle for our future prospects," Nature v.363, 27 (1993).

Gould, S.J., and Eldredge, N., "Punctuated equilibria," Paleobiology v3 No.2, pp. 115-151, (1977).

Hanson, R., "Critiquing the Doomsday Argument," (1998), https://web.archive.org

Kopf, T., Krtous, P., and Page, D., "Too Soon for Doom Gloom?" arXiv 9407002 (1994).

Korb, K.B., and Oliver, J., "A Refutation of the Doomsday Argument," Mind v.107 403-410 (1998).

Leslie, J., "The End of the World," London and New York: Routledge, (1996).

Leslie, J.., "Infinitely Long Afterlives and the Doomsday Argument?," Philosophy Vol.83 No.326 pp.519-524 (2008).

Leslie, J., "Risk that Humans will soon be extinct," Philosophy v85 No.4 (2010).

McCutcheon, R., "What, Precisely, is Carter's Doomsday Argument?" https://philpapers.org, (2018)

McCutcheon, R., "In Favor of Logarithmic Scoring," Phil. Sci. v86 No.32 286-303 (2019)

Monton B. and Roush S., "Gott's Doomsday Argument," https://philpapers.org (2001).

Northcott, R., "A Dilemma for the Doomsday Argument," https://philpapers.org (2016).

Oliver, J. J., and Korb, "A Bayesian Analysis of the Doomsday Argument," Monash U. (1998).

Olum, K., "The Doomsday Argument and the Number of Possible Observers," arXiv gr-qc-0009081 (2001) and Phil.Quarterly v.52 No.207 pp.164-184 (2002).

Sandberg, A., "An Overview of Models of the Technological Singularity," agi-conf.org, (2009).

Sober, E., "An Empirical Critique of Two Versions Of The Doomsday Argument," Synthese v.135 (2003).

Ulam, S., "Tribute to John von Neumann,"" Bull.Amer.Math.Soc. v64 No.3 pp1-49 (1958)

Vinge, V., "The Coming Technological Singularity: How to Survive in the Post-Human Era," NASA Publication CP-10129 (1993).

Wikipedia, https://en.wikipedia.org/wiki/German-tank-problem (2019)