# ANTHROPIC INDEXICAL SAMPLING AND IMPLICATIONS FOR THE DOOMSDAY ARGUMENT

MATT CUSHMAN

ABSTRACT. Anthropic reasoning refers to a class of arguments that incorporate the information entailed by our own existence to make inferences about the world in which we live. One prominent example is the *Doomsday Argument*, which makes predictions about the future total population of human observers yet to be born based on the ordinal rank of our birth among humans that have been born so far. A central question in anthropic reasoning is from which distribution should we consider ourselves to be randomly sampled. The Self Sampling Assumption (SSA) states that we should reason as if we're a random sample from the set of actual existent observers, while the self indication assumption (SIA) states that we should reason as if we're a random sample from among the set of all possible observers (see [1]). Effectively, SIA weighs the probability of our actual world by the number of observers relative to SSA. The distinction is important, as SSA supports the Doomsday Argument, while SIA refutes it. We consider a new thought experiment called *Geometric Incubator* and show that SSA implies precognition of coin flips in this hypothetical world. We consider this to be very strong evidence in favor of SIA over SSA and against the Doomsday Argument. We use this observation to develop a more axiomatic mathematical theory of anthropic reasoning. We also introduce an empirical version of the Doomsday Argument.

## 1. Introduction: SSA versus SIA

Anthropic reasoning is an umbrella term used to encompass philosophical and physical arguments that attempt to draw conclusions about the universe using our own existence as empirical evidence. For example, the *Weak Anthropic Principle* states that the laws of the universe have to be consistent with our own existence, while the *Strong Anthropic Principle* goes further and asserts that the universe is somehow compelled to have conditions that support the development of intelligent life, capable of making observations. The first principle is fairly uncontroversial while the second is the subject of much debate.

A related argument is the *Doomsday Argument* (DA), which states that we should regard our birth rank among all humans who have or will have lived as being randomly sampled from the uniform distribution over such ranks (namely, the set $\{1 \ldots n\}$ where $n$ is the total population of humans). Gott's version ([3]) can be summarized as follows: let $k$ be the number of humans born before us (so that we're the $k+1$-st human. We use this notation as a standard in this paper). Gott's Doomsday Argument would conclude, for example, $n < 20k$ with 95% probability. By modeling the likely future trajectory of human population, we can then infer a high probability upper bound on the end data for all of humanity.

Another version by Leslie ([4], [5]) simply argues that a Bayesian shift in probabilities in favor of earlier demise of humanity is warranted. While the conclusion is somewhat weaker and less specific numerically, Leslie takes considerable care to consider the philosophical assumptions underpinning his argument.

There are two common "assumptions" that have been debated in the philosophical community to address such questions (they are not so much "assumptions", but rather /em proposals on how to properly assign credences to various states of the world). Known as the "self-sampling assumption" or SSA and the "self-indication assumption" or "SIA", the difference focuses on how one should condition on indexical information including knowledge of one's birth rank. These two assumptions are usually stated as follows:

- Self Selection Assumption (SSA): All other things equal, an observer should reason as if they are randomly selected from the set of all actually existent observers (past, present and future) in their reference class.
- Self Selection Assumption (SIA): All other things equal, an observer should reason as if they are randomly selected from the set of all possible observers.

As Olum ([6] points out and we explicate below, the Doomsday Argument holds under SSA, while the probability shift in SIA exactly cancels out Leslie's probability shift in SIA. Hence, under SIA, the Doomsday Argument would be considered invalid. As mentioned above, we argue below that SIA is the only reasonable way to assign credences without reliance on some sort of supernatural capabilities.

A clear distinction between SSA and SIA is demonstrated by the *Incubator* thought experiment ([1]):

**Thought Experiment 1** (Incubator)**.** *In an otherwise empty world, a machine called "the incubator" kicks into action. It starts by tossing a fair coin. If the coin falls tails then it creates one room and a man with a black beard inside it. If the coin falls heads then it creates two rooms, one with a black-bearded man and one with a white-bearded man. As the rooms are completely dark, nobody knows his beard color. Everybody who?s been created is informed about all of the above. You find yourself in one of the rooms.*

The basic question is: what credence do you assign to the coin having had landed heads? In SSA, being born provides no information about the coin flip, as your existence is compatible with both results of the coin flip. Hence, the credence is 0.5.

In SIA, the answer differs. When we awaken, we consider three possibilities: tails and we're the first person, or heads and we're one of two distinct people. Hence, our credences at 1/3 for tails and 2/3 for heads.

What is the essential difference between the two? In both, there is a set of objective, independent of observer "world-states", given by the coin flips. We'll call this set $X = \{H, T\}$, with probability distribution $\mathrm{P}\{F = H\} = \mathrm{P}\{F = T\} = 0.5$. For each world-state, there is a number of observers $N(x)$ for $x \in X$. In other words, $N$ is a $\mathbb{N}$-valued random variable. SSA says that the subjective probability distribution on $X$ of an observer emerging from the incubator remains unchanged, while SIA says the observer should reweigh it by the number of observers created.

More generally and formally, we characterize SSA and SIA in the following manner. Given a probability distribution on the total population of observers $N$, then SSA and SIA are methods of computing subjective credences $\nu(n,k)$ for a member of the population on the set of pairs $(n,k)$, where $n$ is the total populations of observers and $k$ is the number of observers created prior to the creation of that new observer (so that $k = 0$ for the first observer and $k = n-1$ for the final observer). Let $K$ be the random variable of our birth index (the number of humans having been born before us). Both SSA and SIA satisfy the following two properties:

**Definiton 1** (Consistency)**.** *The credence that $K \geq N$ is zero. The observer can't have been created AFTER all observers have been created.*

**Definiton 2** (Copernican Principle)**.** *For any two natural numbers $k_0$ and $k_1$ less than $n$, the credence of $(N,K) = (n,k_0)$ and $(N,K) = (n,k_1)$ should be the same (no one has any special, privileged position* a priori*. We are equally likely to have been created with any observer's birth rank).*

They do differ in the following way, however: (Note: we are writing probability mass functions of the discrete random variables $N$ and $K$ as $\nu(n,k)$ in these equations for notational simplicity)

SSA Self sampling assumption. The marginal distribution of $N$ for a newly created observer is the prior distribution of $N$. So, for $k < n$:

$$\nu_{\text{SSA}}(n,k) = \frac{\mu(n)}{n}$$

SIA Self indication assumption.The marginal credence of $n$ is the prior distribution $n$ weighted by the number of obsevers. This is itself a new probability distribution on $X$, which we denote $\mu_{\text{SIA}}(n)$. So, for $k < n$:

$$\nu_{\text{SIA}}(n,k) = \frac{\mu_{\text{SIA}}(n)}{n}$$

$$\mu_{\text{SIA}}(n) = \frac{n \cdot \mu(n)}{\sum_{m=0}^{\infty} m \cdot \mu(m)}$$

What arguments are there for and against SSA and SIA? There are many, with the most crucial argument against SIA being the "presumptuous philosopher".

**Thought Experiment 2** (Presumptuous Philosopher)**.** *It is the year 2100 and physicists have narrowed down the search for a theory of everything to only two remaining plausible candidate theories, T1 and T2 (using considerations from super-duper symmetry). According to T1 the world is very, very big but finite, and there are a total of a trillion trillion observers in the cosmos. According to T2, the world is very, very, very big but finite, and there are a trillion trillion trillion observers. The super-duper symmetry considerations seem to be roughly indifferent between these two theories. The physicists are planning on carrying out a simple experiment that will falsify one of the theories. Enter the presumptuous philosopher: "Hey guys, it is completely unnecessary for you to do the experiment, because I can already show to you that T2 is about a trillion times more likely to be true than T1 (whereupon the philosopher explains* SIA*)!*

Indeed, it seems magical that the presumptuous philosopher could conjure such a result out of thin air. The methodological supremacy of empirical science, as opposed to armchair philosophizing, seems to dictate that we need to accept SSA over SIA. However, we prove below that the reverse is actually the case: it is SSA that enables paradoxical predictions by armchair philosophizing (a particularly stark example being probabilistic *precognition* of the results of coin flips), and SIA that saves us from it. While it is true that SIA would enable us to draw the conclusion given the narrow and contrived statement of the problem, there is no paradox as discussed below in Section 5.

## 2. Geometric Incubator

We begin with a modification of Thought Experiment 1:

**Thought Experiment 3** (Geometric Incubator)**.** *Suppose in an otherwise empty universe a machine attached to an incubator flips a possibly biased coin which lands heads with probability* $p \in (0, 1)$ *and tails with probability* $1 - p$*. The coin flips are independent events. The machine continues flipping coins until a tails appears so that the sequence of flips is some number – possibly zero – of heads followed by a single tails. After each heads , the machine runs the incubator, creating a new human being.*

Suppose you wake up as an observer in this world and want to make statistical inferences about the future coin flips and the distribution of $n$, the total population of humans. In particular:

(1) What is the probability that the next coin flip is tails?
(2) What is the probability distribution you should assign to (random variable) $N$, the number of human beings that will ever be created in this universe?

Let $X$ be the set of all such sequences of flips. As in Incubator above, we think of $X$ as being the set of "states of the world" for this universe. We can identify $X$ with $\mathbb{N}$ by identifying a sequence of coin flips in $X$ with the number of heads in the sequence.

The one piece of information you gain upon emerging from the incubator is that you are the $K + 1$-th human: you can observe $K$ humans that were born before you for some natural number $K$. The total final population of humans $N$ is simply the number of heads in the completed sequence of flips. By definition, as you observe $K$ other humans, we must have $K < N$. Therefore, the probability of the next coin flip being tails is the same as the probability that you are the final human. In other words, letting $NT$ be the event that the $k + 2$nd (which is the next) coin flip is tails, then $\mathrm{P}(NT) = \mathrm{P}(N = K + 1)$. Hence, questions 1 and 2 are related. *A priori*, $n$ is distributed *geometrically.* The probability of exactly $n$ observers is the same as flipping $n$ heads in a row followed by a tails. The probability mass function of the distribution of $N$ is $\mathrm{P}(N = n) = p^n(p - 1)$ for $n \in \mathbb{N}$. In particular, because $p < 1$ we see $Pr(n = 0) = 1 - p > 0$.

This system is *memoryless* in the sense that knowing that the population count $N$ is at least $n_0$ for some fixed $n_0$ provides no additional information on the distribution of $N - n_0$. More precisely, the distribution of $N - n_0$, conditioned on $N \geq n_0$ is also geometrically distributed with parameter $p$. Intuitively, knowing the first $n$ flips tells us nothing about the subsequent flips. The distribution of the number of humans yet to be created after the creation of the $n_0$-th human is independent of $n_0$.

Thus, the following fact seems self-evident: *Each human created should believe that the probability that the next coin flip is tails is $1-p$, regardless of their birth rank $k$, the number of observed humans so far, and whether or not they've see the rest of the human population or not.* We refer to this property as the *no-precognition property.* In fact, given the assumptions of the thought experiment, no human should have any information about any future coin flip, and should continue to have credence $p$ for it coming up heads. The negation would be a form of *precognition*: at least one human being would be created with some probabilistic information about the result of some future coin flip.

We now apply the definitions of SSA and SIA to determine the subjective probability distribution of the number of future humans for each human as they are created. As mentioned above, we will see that in the case of SSA, the probability of tails will actually exceed $1-p$. This is the famous *Doomsday Argument* applied to our toy universe. Under SIA, the probability remains unchanged.

2.1. **The SSA Case.** Recall that under SSA, the probabilities of the underlying state space are unchanged and the various indices within the population count $n$ are equally probable. Let $\nu_{\mathrm{SSA}}$ be the probability measure on pairs $(n,k) \in BBN \times \mathbb{N}$ corresponding to SSA. For convenience, we will slightly abuse this notation and write $\nu_{\mathrm{SSA}}(n,k)$ for the credence that a human would have for the statement "I am the $k+1$ human born out of a total population of $n$ humans" (the point density function for this distribution). According to the definition of SSA, this is the unique probability distribution $\nu_{\mathrm{SSA}}$ on pair of natural numbers $(n,k)$ such that

(1) $\mathrm{P}\{k \geq n\} = 0$
(2) Each index $k$ compatible with a given population size $n$ has the same credence. In other words, $\nu_{\mathrm{SSA}}(n,k_0) = \nu_{\mathrm{SSA}}(n,k_1)$ for $k_0, k_1 < n$.
(3) The marginal probability distribution of $n$ under $\nu_{\mathrm{SSA}}$ matches that of $\mu$

Translating directly into an explicit formula for $\nu_{\mathrm{SSA}}(n,k)$ is straightforward, just dividing the pdf of the geometric distribution on $n$ equally across the $n$ values of $k$ that are non-zero:

$$\nu_{\mathrm{SSA}}(n,k) \quad = \quad \begin{cases} 0 & \text{if } k \geq n \\ \frac{p^{n-1}(1-p)}{n} & \text{if } k < n \end{cases}$$

One minor technical note: it would be tempting to write $p^n(1-p)/n$ for the $k < n$ case above, however this is not a probability distribution because there is no $k < n$ in case $n = 0$. So, we need to normalize by the set of cases in which $n > 0$, which happens precisely when the first flip is heads, i.e. $p$ of the time. This is related to the "dicontinuity" noted by Olum and referred to previously.

2.2. **The SIA Case.** We obtain different answers under SIA of course. We need to weight the probability of each sequence of coin flips by the number of observers produced in each.

If $n+1$ is the total number of coin flips, there are $n$ heads followed by a single tail, which results in $n$ observers. This is a slightly more involved "global" computation but still easily done. The probability distribution on $n$ is proportional to $n(1-p)p^n$. This needs to be normalized to have sum 1 over all $n$. Using basic calculus (recounted in Lemma 1 at the end of this paper) we obtain

$$\begin{aligned}\sum_{n=0}^{\infty} np^n(1-p) &= (1-p)\frac{p}{(1-p)^2} \\ &= \frac{p}{1-p}\end{aligned}$$

So, $\mu_{\text{SIA}}(n) = n(1-p)p^n/(p/(1-p)) = np^{n-1}(1-p)^2$. Hence, given $k, n \in \mathbb{N}$ with $k < n$, the probability (under SIA) that we're the $k+1$-th individual of $n$ total observers is $\nu_{\text{SIA}}(n,k) = \mu_{\text{SIA}}(n)/n = p^{n-1}(1-p)^2$. Putting it all together, the point density function for SIA is:

$$\nu_{\text{SIA}}(n,k) = \begin{cases} 0 & \text{if } k \geq n \\ p^{n-1}(1-p)^2 & \text{if } k < n \end{cases}$$

2.3. **Coin Flip Credences.** Let us return to our original questions. Suppose we emerge from the incubator an observer exactly $k$ other humans that were created before us (so that we're the $k+1$ human created). What is the probability that the next flip is tails? This is equivalent to the probability that there are exactly $n = k+1$ humans given that there are at least $k+1$ humans. In terms of our indexical space, $\nu(k+1,k)/\sum_{n=k+1}^{\infty}\nu(n,k)$.

Under the assumptions of SSA we can compute this as follows:

$$\begin{aligned}\mathrm{P}_{\text{SSA}}(\text{next tails}|\text{created as } k+1\text{-th person}) &= \frac{\nu_{\text{SSA}}(k+1,k)}{\sum_{n=k+1}^{\infty}\nu_{\text{SSA}}(n,k)} \\ &= \frac{p^k(1-p)/(k+1)}{\sum_{n=k+1}^{\infty}p^{n-1}(1-p)/n} \\ &= \frac{p^{k+1}/(k+1)}{\sum_{n=k+1}^{\infty}p^n/n} \\ &\gneq \frac{p^{k+1}/(k+1)}{\sum_{n=k+1}^{\infty}p^n/(k+1)} \\ &= \frac{1}{\sum_{m=0}^{\infty}1/p^m} \\ &= 1-p\end{aligned}$$

So, in general, this is greater than and not equal to $1-p$. Thus SSA in this case enabled *precognition* of the future coin flip. The probability of tails is strictly larger than the intrinsic probability $1-p$ of tails. Intuitively this does make sense given the stated logic behind SSA. If we assume that we're a random sample among the set of humans that actually exist, our index in this set does imply something about the total set of humans to exist in the future, and hence the probability of the coin flip.

Conversely, for the SIA case we see

$$\begin{aligned}
\mathrm{P}_{\mathrm{SIA}}(\text{next tails}|\text{created as } k+1\text{-th person}) &= \frac{\nu_{\mathrm{SIA}}(k+1,k)}{\sum_{n=k+1}^{\infty}\nu_{\mathrm{SIA}}(n,k)} \\
&= \frac{p^k(1-p^2)}{\sum_{n=k+1}^{\infty}p^{n-1}(1-p^2)} \\
&= \frac{p^{k+1}}{\sum_{n=k+1}^{\infty}p^k} \\
&= \frac{1}{\sum_{m=0}^{\infty}p^m} \\
&= (1-p)
\end{aligned}$$

This is exactly the probability one would expect to see tails. Thus, there is no precognition in this case.

To summarize: under the assumptions of SSA, the probability of tails for the next flip for the $n$-th observer born is strictly greater than $1-p$. However, for SIA, the probability remains precisely $1-p$. This suggests a deeply paradoxical, borderline paranormal, result regarding SSA which we therefore reject.

This argument is in fact a mathematical generalization of the "discontinuity" argument against SSA mentioned in [6] (page number ?). Olum points out that in SSA, every world with an observer consistent with our existence receives equal weight, while of course worlds with no observers receive no weight. So, for example, the standard Incubator problem gives probability 1/2 to both heads and tails as in both cases the world contains observers whose experience is consistent with our experience. However, if one modifies Incubator to give either 0 or 1 observer, we know the result of the coin flip with certainty. This discontinuity seems dissatisfying in this particular case. In a sense, our argument shows that this discontinuity in the extremes actually is a result of consistent precognition effects that pervade many more regular cases.

A related thought experiment was proposed by Leslie himself in [5]. Called the *Shooting-Room Paradox*, it proceeds as follows:

**Thought Experiment 4** (Shooting-Room)**.** *First a batch of ten people are led into this room. A pair of dice is thrown in front of their eyes. If a double six comes up they are all shot. Otherwise they leave the room safely and a new batch, this one containing a hundred people, is thrust in. The process continues, with each consecutive batch ten times larger than the previous one, until there is a double six; whereupon the people in the room at that time are shot and the experiment ends. (quoted from Bostrom).*

Leslie's observation is the paradox that, since the probability that any one generation is killed is 1/36, this should be the probability of survival. However, after the fact on can easily compute that approximately 90% of the people that are ever in the room are shot. By this measure, it looks like the probability should be 90%.

There have been a number of papers addressing this paradox (for example **??**). Leslie focuses on the appartent paradox caused by the disparity between the proportion of people even in the room that are shot, versus the probability that a specific generation will be shot virtue of the next die roll. This paradox has been thoroughly analyzed elsewhere, and indeed one could reason analogously to the above analysis of Thought Experiment 3 (Geometric Incubator)to land on the correct result of 1/36. Indeed, the number of batches in the room follows a geometric distribution, although the total number in room is different because of the exponential growth of each batch size. To our knowledge, no one has used this to observe that SIA would lead to precognition of the dice rolls, however.

## 3. Three versions of Skynet and Occam's Razor

The Doomsday Argument can be compared to the German Tank Problem. In World War II, the Allied forces wanted to estimate the production capacity of the Wehrmacht – in particular how many tanks the Germans were producing. There were various methods employed, such as traditional spying and reconnaissance. One of the most accurate was a purely statistical method: analyzing the serial numbers on parts in captured German tanks. The basic argument is that, if you capture a single randomly chosen tank and the serial number on the tank is $n$, then one could estimate the total population of tanks (which is roughly $2n$). There are easy generalizations to the case of $k > 1$ tanks having been captured.

The success of this method for estimating various quantities could be taken as a strong argument in favor of SSA and the Doomsday argument. The underlying arguments in both cases are similar. However, we argue that they are actually fundamentally different, as the following thought experiment will show. We take this as another argument in favor of SIA over SSA.

**Thought Experiment 5** (Skynet)**.** *As in the Terminator film franchise, artificially intelligent robots (Terminators) are attempting to exterminate humanity, led by a central AI Skynet. They are opposed by John Connor, a human. Skynet is capable of creating Terminators using a machine, but the machine that creates them will irrevocably crash with probability* $1 - p$ *after right before each run (so the probability of a successful run it* $p$*). Everybody knows the value of* $p$ *(both the Terminators and their human foes). The machine has run and generated a number* $N$ *of Terminators. The machine stamps each Terminator with a sequential serial number. Consider the three scenarios:*

1. *You are John Connor, and are told by a spy that there exists a Terminator with serial number* $K$*.*
2. *You are John Connor, and capture a Terminator uniformly sampled from among the population of all Terminators. Its serial number is* $K$*.*
3. *You awaken as a Terminator, and observe you have serial number* $K$*.*

*In each case, what are the credences you should assign to the total population of Terminators?*

Item 3 is really the case of interest: as a Terminator, what should your credences be conditioned on your index? This is analogous to being born a human being and aware of your birth rank being $k$. The first two are shed light on the third, however.

In the first case, an outside observer (John Connor) learns only that there are at least $K$ Terminators. There is no other information or context provided. The Terminators are generated by a process that results in a geometric distribution, and again using the memoryless property of such distributions, the distribution of the number of Terminators created after the $K$-th, conditioned on there being at least $K$, is again geometric with the same parameter.

In the second case, the usual arguments from the German tank problem pertain, and we gain more information about the resulting distribution. If we randomly manage to capture the Terminator with serial number $10^9$, it seems a lot less likely that there are only 2 more Terminators than if we capture the Terminator with serial number 5. The frequentist approach yields a point estimate of the total population of Terminators of $2K$, with a 95% confidence interval of $[K, 20K]$. The Bayesian analysis does not yield a convergent mean, but the distribution can be computed.

Note that capturing a uniformly distributed Terminator with serial number $K$ entails that there are at least $K$ Terminators. So, more information is revealed to John Connor in Item 2 than in Item 1. In general: knowing that a random variable $N$ exceeds a particular variable $K$ is not the same as knowing a sampled value $K$ from anuniform distribution between 0 and $N$. The latter strictly entails the former.

Which one is most pertinent for the real case of interest, Item 3? The mathematical argument described in the previous section for Geometric Incubator still pertains, which favors Item 1 over Item 2. Claiming Item 2 is the right model for Item 3 would imply some sort of precognition on the part of the Terminator.

A more universal principle is at work. In the comparison of Item 1 versus Item 2, there is strictly more information revealed to John Connor. The information revealed to the Terminator in Item 3 is compatible with both Item 1 and Item 2, but it's not at all clear why the Terminator would be justified in thinking of himself as having been sampled uniformly as in the case of Item 2. It is certainly the case that we, as part of the Copernican principle, think of ourselves as having each birth rank being equally likely. While this does seems to suggest that we should regard our birth rank as having actually been sampled uniformly from among the observers created, this is an illusion. Each birth rank being equally likely is compatible with both SIA as well as SSA. Yet only SSA entails the additional information, of the sort in Item 2, over and above SIA which is analogous to Item 1. It's not at all evident where this additional information would come from, from the mere fact of having been an observer created with a particular birth rank. Therefore, an argument for the most parsimonious model – "Occam's Razor" – would favor SIA over SSA.

## 4. The General Case

We have seen that in the particular case of 3 that there is precognition under the assumptions of SSA and no precognition in the case of SIA. In this section we show that, in general, SIA is the only assignment of probabilities to the sampling space that eliminates precognition.

4.1. **Framework.** We begin with a probability space $(X, \mathcal{F}, \mu)$. We think of $X$ as the set of possible *states of the world*, with probabilities given by $\mu$. To each state of the world $x \in X$, there is a corresponding *total population count* $N(x) \in \mathbb{N}$. $N$ is a random variable on $X$ with values in the natural numbers $\mathbb{N}$, equal to the total number of observers created in the world-state $x$. Note, $N(x)$ may be zero for some $x$, although we do require it to be finite for almost all $x$. For some results below, will need to assume that $N$ has finite expected value as well.

This data, $(X, \mathcal{F}, \mu, N)$, is meant to capture the "objective state" of the universe. By objective state, we mean the observations about the universe that are independent of any particular observer. An observer outside of the system should have credences and be able to reason about these states. We will refer to the observers inside the system as "internal observers" and hypothetical observers outside the system as "external observers". For example, in Incubator, the humans being incubated are internal observers and we (or the machine) would be external observers. External observers share the credences given by $\mu$. These external observers might not actually exist in any sense, but rather be a convenient fiction.

Observers within the system (internal observers) should also be able to reason about the objective state of the universe (in other words, assign credences to events such as $x \in S$ for some $S \in \mathcal{F}$). The states make no reference to any observer specific indexical information (for example, statements like "I am the $k$-th observer created" cannot be referenced by subset of $X$) and any two rational observers with identical information (whether in or outside the system) would have the same credences on $X$.

We now introduce a formalism that does allow for reference to internal-observer specific indexical information. The *indexical sampling space* for $X$ is $Y = \{(x, k) \in X \times \mathbb{N} | k < N(x)\}$. Intuitively, an element of $Y$ corresponds to a particular internal observer in a particular possible world. The element $(x, k) \in Y$ is thought of as referring to the $(k+1)$-st observer created in the world corresponding to $x$, in which there are $N(x)$ observers, and we will refer to these elements as *observer-states*. As before, the index $k$ runs from $0$ to $N(x) - 1$ for notational convenience. In contrast to $X$, an element of $Y$ picks out a particular observer and hence is subjective. Two internal observers would have different credences for subsets of $Y$, precisely because the true, underlying state corresponding to those two observers is actually different (they share the same first coordinate in $X$, but two different values in $\{0, \dots, N(x) - 1\}$ for the second coordinate).

Let $\pi_X : Y \to X$ be the projection onto the first element. Then $N \circ \pi_X$ is a random variable from $Y$ to $\mathbb{N}$, which we will also write as simply $N$. The projection onto the second coordinate $K : Y \to \mathbb{N}$ is also a $\mathbb{N}$-valued random variable. Formally, $N(x, k) = N(x)$ and $K(n, k) = k$. As mentioned above, $0 \le K < N$ everywhere on $Y$.

In a sense, $\pi_X$ and $K$ perform complementary functions. The first, $\pi_X$, "forgets" all observer-dependent structure and retains only the observer-independent state. Conversely, $K$ only retains the indexical information corresponding to an observer.

We will want to consider events on $X$ as events on $Y$ as well. This will enable us to translate between external and internal observers. Internal observers can refer to certain events, namely ones that include indexical information such as "I am observer number 10 created", but every event that an external observer can refer to also makes sense from the perspective of an internal observer. We do this simply by taking the pre-image of the event under the map $\pi_X$. For convenience we notate this by $\tilde{S} = \pi_X^{-1}(S)$ for $S \in \mathcal{F}$.

We also define subsets of $X$ and $Y$ by the number and index of observers. The first, $X_n$ is the set of world-states with $n$ observers, and $Y_{n,k}$ is the set of $k$-th indexed observer-states in worlds with $n$ observers.

$$\begin{aligned} X_n &:= N^{-1}(n) \\ &= \{x \in X | N(x) = n\} \\ Y_{k,n} &:= \tilde{X_n} \cap K^{-1}(k) \\ &= \{(x, j) \in Y | N(x) = n \text{ and } k = j\} \end{aligned}$$

We want to characterize probability distributions $\nu$ on $Y$ that capture the intuitive notion of anthropic sampling. The measure $\nu$ would be defined on the $\sigma$-algebra $\mathcal{G}$ given by the restriction of the product $\sigma$-algebra on $X \times \mathbb{N}$, where $\mathbb{N}$ has the discrete $\sigma$-algebra (i.e. the power set $\sigma$-algebra, in which every subset of $\mathbb{N}$ is measurable). Both projections $\pi_X$ and $K$ are measurable, hence $K$ and $N$ are random variables on $Y$.

The probability measure $\nu$ should capture the credences of an event (i.e. subset $S \subset Y$) that a rational observer would have, knowing the probability distribution $\mu$ on $X$ and the total population count $N$.

**Definiton 3** (Anthropic Sampling Distrubition $\nu$)**.** *For any (measurable) subset $S$ of $Y$ (i.e. $S \in \mathcal{G}$), $\nu(S)$ is the credence that an rational internal observer give to the statement that "The world is in state $x \in X$ and I am observer number $k \in \mathbb{N}$ for some $(x, k) \in S$", immediately after being created, knowing the probability distribution of $\mu$ on $X$ and population count random variable $N$, but prior to any other knowledge being imparted.*

The fundamental problem of anthropic reasoning in our language is: starting with $(X, \mathcal{F}, \mu, N)$ as above, consider the indexical sampling space $Y$ a defined above,hat is the most appropriate probability measure $\nu$ to put on $Y$ to reflect the credences of an observer $y$ who finds him/herself in a world $x$ sampled from $X$ according to distribution $\mu$ with population $N(x)$. Is this even a well-defined, unique probability distribution?

Merely being a probability distribution on $Y$ puts some constraints on $\nu$, but more is needed to fully characterize and capture the notion of Anthropic sampling. The Copernican Principle as defined above (Defintion 2) is one obvious constraint. Restating it formally for this more general framework:

**Definiton 4** (Copernican Principle)**.** *For all $k_0 < n_0 \in \mathbb{N}$ and $S \in \mathcal{F}$ with $N(S) = n_0$ ($N$ is constant on $S$ with value $n_0$), $\nu(K = k_0 | \tilde{S}) = 1/n_0$.*

In other words: knowing that there are precisely $n_0$ observers (and perhaps some additional information that refers only to the objective state of the world but nothing referring to our birth rank), then our birth rank is distributed uniformly between 0 and $n_0 - 1$.

The Self-Sampling Assumption and Self-Indication Assumption that we discussed earlier both purport to answer this question. We translate them into this new formalism as follows:

**Definiton 5** (Formal Self-Sampling Assumption)**.** *Let $\nu$ be a measure on $(Y, \mathcal{G})$ that satisfies the Copernican Principle. The (Formal) Self-Sampling Assumption for $\nu$ states that the $\nu$-measure of a set of world-states is the same as the $\mu$-measure, conditioned on there being some observers at all:*

$$\nu_{\text{SSA}}(\tilde{S}) \quad = \quad \mu(S|N > 0)$$

**Definiton 6** (Formal Self-Indication Assumption)**.** *Let $\nu$ be a measure on $(Y, \mathcal{G})$ that satisfies the Copernican Principle. The (Formal) Self-Indication Assumption states that the $\nu$-measure of a set of world-states is the $\mu$-measure weighted by the number of observers in that state, normalized to be a probability measure by the expected number of observers:*

$$\nu_{\text{SIA}}(\tilde{S}) \quad = \quad \frac{\int_S N(x) d\mu(x)}{\int_X N(x) d\mu(x)}$$

Note that the normalization factor needed in the definition of $\nu_{\text{SIA}}$ to make it a probability measure is simply the expected number of observers, which we assume is finite.

We will posit another constraint on $\nu$ that, combined with the Copernican Principle, will fully characterize $\nu$. Arguing in a similar manner as in the case of Geometric Incubator, that the "Self-indication assumption" is the only assignment of credences on $Y$ that satisfies this both the Copernican Principle and this additional property.

What is missing beyond the Copernican Principle and the construction of $Y$ itself (which guarantees consistency and that $N$ and $K$ are random variables – are measurable – on $Y$) is some condition linking $\mu$ (a measure on $X$) and $\nu$ (a measure on $Y$). The credence of an internal observer ought to be linked to the credence of an external observer when they share "similar information". How can we formalize this intuition? What does it mean for an internal and external observer to have "similar" information?

We formalize this notion as follows: Internal observers should have credences identical to those an external observer would have who learns that the given internal observer has been created (or at least one with that same birth rank has been created). This can be translated into the following condition relating $\mu$ and $\nu$:

**Definiton 7** (No-Precognition Principle)**.** *For every event $S \in \mathcal{F}$ (i.e. referring only to external world-states in $X$) and natural number $k$, the credence that an internal observer gives to $\tilde{S}$ knowing only that they are observer number $k+1$ is the same as the credence that an external observer gives to $S$ knowing only that there are more than $k$ observers. Formally: for all $S \in \mathcal{F}$ and $k_0 \in \mathbb{N}$, $\nu(\tilde{S}|K = k_0) = \mu(S|N > k_0)$.*

Why do we refer to this as "no-precognition"? Because if these credences do not agree, then the internal observer gains some additional knowledge about the objective state of the world (reflected in the change in probability measure) beyond that of the external observer merely by virtue of the fact that THEY (as opposed to someone else) is the $k$-th internal observer created. This could be literal "pre-cognition" in the sense of foreknowledge, or (if we regard the events as having been determined already) some sort of "extra-sensory perception". Either way, the internal observer has some unexplained, non-natural insight into which $x \in X$ corresponds to their universe. This is merely a generalization of the observation we made before in Geometric Incubator, where the subjective credence of the next coin flip for the internal observer differs from the credence assigned by the objective observer.

**Theorem 1.** *Let $(X, \mathcal{F}, \mu)$ be a probability space and $N : X \to \mathbb{N}$ be a random variable as described above. Let $\lambda$ be the expected population count: $\lambda = \mathbb{E}_X[N]$. Construct the indexical sampling space and associated $\sigma$-algebra $\mathcal{G}$ as above. Consider a probability measure $\nu$ on $(Y, \mathcal{G})$. The following are equivalent:*

(1) *Both the No-Precognition Principle (Definition 7) and Copernican Principle (Definition 4) hold for $\nu$.*
(2) *For all $S \in \mathcal{F}$ and $k_0, n_0 \in \mathbb{N}$ with $N(S) = \{n_0\} > k_0$, $\nu(S \times \{k_0\}) = \mu(S)/\lambda$*
(3) *For all $T \in \mathcal{G}$, $\nu(T) = \sum_{k=0}^{\infty} \mu(\{x \in X | (x, k) \in T\})/\lambda$.*
(4) *$\nu$ is the product measure on $X \times \mathbb{N}$ (using the counting measure on $\mathbb{N}$) restricted to $Y$ and normalized to have total measure 1 (in other words, to be a probability measure). The normalizing factor is $1/\lambda$.*
(5) *$\nu$ satisfies the Self-Indication Assumption (Definiton 6).*

*Proof.* The hardest part is to see (1) $\iff$ (2). Assume that $\nu$ satisfies the Copernican principle, as that is part of the assumption in both (1) and (2).

Let $S, n_0, k_0$ be given as in (ii) and assume that the No-Precognition Condition holds for $\nu$. We expand both sides of the "no-precogniton" equation. First the left hand side, using the Copernican principle at the second to last equation:

$$\begin{aligned}
\nu(\tilde{S}|K = k_0) &= \frac{\nu(\tilde{S} \wedge (K = k_0))}{\nu(K = k_0)} \\
&= \frac{\nu(S \times \{k_0\})}{\sum_{n=0}^{\infty} \nu((K = k_0) \wedge (N = n))} \\
&= \frac{\nu(S \times \{k_0\})}{\sum_{n=0}^{\infty} \nu(K = k_0 | N = n) \nu(N = n)} \\
&= \frac{\nu(S \times \{k_0\})}{\sum_{n=k_0+1}^{\infty} \nu(N = n)/n}
\end{aligned}$$

Similarly, for the right hand side we obtain:

$$\mu(S|N > k_0) = \frac{\mu(S)}{\sum_{n=k_0+1}^{\infty} \mu(N = n)}$$

Equating both expressions, in order to show (2) it will suffice to prove that:

$$\sum_{n=k_0+1}^{\infty} \mu(N=n) = \mathbb{E}_X[N] \cdot \sum_{n=k_0+1}^{\infty} \nu(N=n)/n$$

holds assuming the No-Precognition condition holds.

To see that, we apply the no-precognition property again, this time to $S = \{N = k_0+1\}$:

$$\begin{aligned}\nu(N = k_0+1|K = k_0) &= \mu(N = k_0+1|N > k_0) \\ &= \frac{\mu((N = k_0+1) \wedge (N > k_0))}{\mu(N > k_0)} \\ &= \frac{\mu(N = k_0+1)}{\sum_{n=k_0+1}^{\infty} \mu(N = n)}\end{aligned}$$

which is also:

$$\begin{aligned}\nu(N = k_0+1|K = k_0) &= \frac{\nu((N = k_0+1) \wedge (K = k_0))}{\nu(K = k_0)} \\ &= \frac{\nu((N = k_0+1) \wedge (K = k_0))}{\sum_{n=k_0+1}^{\infty} \nu((N = n) \wedge (K = k_0))} \\ &= \frac{\nu(N = k_0+1)/(k_0+1)}{\sum_{n=k_0+1}^{\infty} \nu(N = n)/n}\end{aligned}$$

Hence:

$$\frac{\nu(N = k_0+1)/(k_0+1)}{\sum_{n=k_0+1}^{\infty} \nu(N = n)/n} = \frac{\mu(N = k_0+1)}{\sum_{n=k_0+1}^{\infty} \mu(N = n)}$$

Applying Lemma 2 (with $a_n = \nu(N = n)/n$ and $b_n = \mu(N = n)$) we conclude $\nu(N = k_0+1)/(k_0+1) = \gamma \cdot \mu(N = k_0+1)$ for all $k_0$ for some $\gamma$. In other words, $\nu(N = n) = \gamma n \mu(N = n)$ for all $n \in \mathbb{N}$ (we get the $n = 0$ case because obviously $\nu(N = 0)$ is 0 – there are no observers in worlds with $N = 0$, therefore the probability that an observer is in such a world is 0). Summing and taking the ratio, we see

$$\begin{aligned}\gamma &= \frac{\sum_{n=0}^{\infty} \nu(N = n)}{\sum_{n=0}^{\infty} n\mu(N = n)} \\ &= 1/\mathbb{E}_X[N]\end{aligned}$$

Hence (2) holds.

To see that (2) implies (1), we can essentially just work the same computation in reverse. Assume (2). Using the same expansion of the left and right hand sides of the No-Precognition equation as before:

$$\mu(S|N > k_0) = \frac{\mu(S)}{\sum_{n=k_0+1}^{\infty} \mu(N = n)}$$
$$\nu(\tilde{S}|K = k_0) = \frac{\nu(S \times \{k_0\})}{\sum_{n=k_0+1}^{\infty} \nu(N = n)/n}$$

Thus, it suffices to prove that

$$\nu(S \times \{k_0\}) = \frac{\sum_{n=k_0+1}^{\infty} \nu(N = n)/n}{\sum_{n=k_0+1}^{\infty} \mu(N = n)} \mu(S)$$

Applying (2) this reduces to

$$\lambda = \frac{\sum_{n=k_0+1}^{\infty} \nu(N = n)/n}{\sum_{n=k_0+1}^{\infty} \mu(N = n)}$$

To see this, we apply (2) again to the event $\{N = n\}$ so that for each $k < n$ we have $\nu((N = n) \wedge (K = k)) = \mu(N = n)/\lambda$. Adding up the $n$ disjoint cases as $k$ runs over $\{0, \ldots, n-1\}$ we obtain $\nu(N = n) = n\mu(N = n)/\lambda$. Diving both sides by $n$ and summing over $n > k_0$ we obtain our result.

(2) $\iff$ (3): To see $\implies$, decompose $T \in \mathcal{G}$ into slices of the form $T_{n,k} = T \cap Y_{n,k}$. Using (2) we compute the $\nu$-measure of slice and the sum using countable additivity. For the other direction, (2) is just a special case of (3).

(3) $\iff$ (4): This is the definition of product measure.

(3) $\iff$ (5): This is the definition of the Self-Indication Assumption.

□

## 5. The Presumptuous Philosopher Redux

We've seen that SSA entails precognition in the case of the geometric incubator, while SIA does not. However, essentially a similar criticism has been leveled against SIA in the past: the presumptuous philosopher paradox referred to above (Thought Experiment 2). By weighing the probability of possible world-states by the number of internal observers, it seems to create information from nothing. For instance, wouldn't SIA force us to believe our universe has infinity many observers?

How to resolve this apparent paradox? We would argue the issue is with an inherent ambiguity in the statement of the problem. Consider the following events and specified conditional probabilities from the statement of the Presumptuous Philosopher:

$$
\begin{aligned}
T1 &= \text{There are a total of a trillion trillion } (10^{24}) \text{ observers} \\
T2 &= \text{There are a total of a trillion trillion trillion } (10^{36}) \text{ observers} \\
ME &= \text{I exist, and am observer number 60 billion} \\
E &= \text{All of experimental and theoretical physics considered by the scientist} \\
pr\{T1|E\} &= 0.5 \\
pr\{T2|E\} &= 0.5
\end{aligned}
$$

Translated it our language: $X = T1 \cup T2$, $N(T1) = 10^{24}$, $N(T2) = 10^{36}$ (meaning that $N$ is constant on each of these subsets of $X$), and $K = 60 \cdot 10^9$.

There are multiple interpretations of the probabilities given in the problem statement. Do we consider these events as being events on some external probability space like $(X, \mu)$ in our framework above, or should they be considered as events on the indexical sampling space $Y$ equipped with the only $\nu$ that fits our No-Precognition assumption? Essentially, does the physics $E$ that determines that the two outcomes are equally likely already account for our observation $ME$, or does it not?

In the first case, in which the evidence $E$ does include the observation that we are among the first trillion trillion observers, further conditioning on this fact yields nothing. So, there is no Nobel prize awaiting the philosopher and the probabilities remain the same.

In the second case, with $E$ regarded as an event in the underlying, objective space $X$, we do in fact get a shift of probabilities. The computation exactly matches Bostrom's original computation.

However, this is assuming that the physicists somehow computed some completely objective *a priori* probability that assigns equal weights to these very different populations of observers. Recall the proper interpretation of the indexical information contained in $ME$ from Definition 3 above. The evidence in $ME$ is the most concrete, immediate evidence we have available to us. However, the second case asks us to discard this, and is derived as if we are standing apart from the system. Our thesis is that all reasoning must occur as probabilities on $Y$, not on $X$. Otherwise, we will be vulnerable to belief in precognition as described above.

Another way of thinking of the same issue: suppose we were given $E$ but not our birth rank $K$. Assume that, in the absence of knowing our birth rank, the evidence $E$ is still such that we consider $T1$ and $T2$ equally credible. Then, through advancements in the fields of anthropology and archeology, we discover $ME$, that our birth rank is $60 \cdot 10^9$. In this case, it is SSA that would dictate that the probabilities should shift and more heavily weight $T1$ (this is simply The Doomsday Argument). It seems just as unlikely for this shift to occur, as the prior one.

## 6. Empirical Doomsday

We have argued above in favor of SIA over SSA. Does this entirely eliminate the Doomsday Argument? While it does appear to eliminate the Carter-Leslie (and Gott) style

Doomsday Argument, an empirically grounded Doomsday Argument can be made in this framework. This new Doomsday Argument is intuitively plausible, as it is based on specific observational evidence about the world and population of observers. The specifics rely on a number of modeling assumptions about statistical distributions of life in the universe. While the true distributions are presumably analytically intractable and impossible to estimate reliably, we can glean some general observations using toy models.

The fundamental idea is that each class of observer would have a total population throughout its history sampled from some probability distribution, which itself unknown. By observing one's population rank, and observer can make inferences about this unknown parameter, and hence the distribution of the total population.

For example, consider a model of the total population distributed according to a geometric distribution (again, just as in Geometric Incubator), $n$ Geom($p$) for some $p \in (0, 1)$. The parameter $p$ itself is unknown, but is distributed according to some other probability distribution. Then, being born at some rank $k$ in the history of the life form – in other works, observing that $n \geq k$ – enables us to make some inferences about $p$. Then, these inferences in turn enable further inferences about the ultimate size of the population $n$.

For example, suppose $p$ is distributed uniformly in the interval $(0, 1)$ and we observe our birth rank is $k$. From the definition of geometric distribution:

$$\mathrm{P}\left[n \geq k | p\right] = p^k \tag{1}$$

The unconditional probability that $n \geq k$ is then

$$\int_0^1 p^k dp = \left. \frac{p^{k+1}}{(k+1)} \right|_{p=0}^{1} = 1/(k+1)$$

Thus, if we let $f(p|n \geq k)$ be the pdf of $p$ after making this observation, and $f(p) = \mathbb{1}_{(0,1)}(p)$ the pdf of the uniform distribution for $p \in (0, 1)$, we find:

$$f(p|n \geq k) = \frac{\mathrm{P}\left[n \geq k|p\right]}{\mathrm{P}\left[n \geq k\right]} f(p) \tag{2}$$

$$= (k+1)p^k \tag{3}$$

Thus, the probability that we will see at least $m$ more members of our population is

$$\mathrm{P}\left[n \geq k + m | n \geq k\right] = \int_0^1 p^m (k+1) p^k dp \tag{4}$$

$$= \frac{k+1}{m+k+1} \tag{5}$$

As $k \to \infty$, this probability approaches 1. Hence, if we've already observed a very large population, we may then suppose that it's likely that a large number of future population members are coming.

As an aside, we should note that in this model the total population has infinite mean. We have:

$$\begin{aligned}\mathbb{E}[n] &= \int_0^1 \mathbb{E}[n|p]\,dp \\ &= \int_0^1 \frac{p}{1-p}dp \\ &= \int_0^1 \frac{1}{1-p} - \frac{1-p}{1-p}dp \\ &= \ln(1-p) - 1|_0^1\end{aligned}$$

which does not converge. This peculiar property of this particular distribution is not important for the argument, and many distributions with finite moments would work just as well. For example, the same results hold if we instead model $p$ as being sampled uniformly from the interval $(0, q)$ for some fixed $q < 1$.

## 7. Mathematical Lemmas

**Lemma 1.** *Let* $p \in (0,1)$. *Then*

$$\begin{aligned}\sum_{k=0}^{\infty} p^k &= \frac{1}{1-p} \\ \sum_{k=n}^{\infty} p^k &= \frac{p^n}{1-p} \\ \sum_{k=n}^{\infty} kp^k &= \frac{n(1-p)p^n + p^{n+1}}{(1-p)^2}\end{aligned}$$

*Proof.* The first two are simply geometric sequences. For the third, let $x = \sum_{k=n}^{\infty} kp^k$.

$$\begin{aligned}(1-p)x &= \sum_{k=n}^{\infty} kp^k - \sum_{k=n}^{\infty} kp^{k+1} \\ &= np^n + \sum_{k=n+1}^{\infty} kp^k - \sum_{k=n+1}^{\infty} (k-1)p^k \\ &= np^n + \sum_{k=n+1}^{\infty} p^k \\ &= np^n + \frac{p^{n+1}}{1-p} \\ &= \frac{n(1-p)p^n + p^{n+1}}{1-p}\end{aligned}$$

using 6 above. Diving both sides by $1-p$ yields

$$x = \frac{n(1-p)p^n + p^{n+1}}{(1-p)^2}$$

□

**Lemma 2.** *Let $a_n, b_n$ for $n \in \mathbb{N}$ be sequences such that $a_n \geq 0$ and $b_n \geq 0$ for all $n \in \mathbb{N}$ and with convergent sums. Assume further that for every $n$*

$$a_n \left( \sum_{m=n}^{\infty} b_m \right) = b_n \left( \sum_{m=n}^{\infty} a_m \right)$$

*Then the two sequences are proportional: $(\sum_{m=0}^{\infty} b_m) \cdot a_n = (\sum_{m=0}^{\infty} a_m) \cdot b_n$ for all $n$. In particular if either sequence is not always zero, then the other is also not always zero and $a_n = \lambda b_n$ where*

$$\lambda = \frac{\sum_{n=0}^{\infty} a_n}{\sum_{n=0}^{\infty} b_n}$$

*The converse is also true: if the sequences are proportional, then the equality holds for all $n$.*

*Proof.* We use induction on $n$. Let $n \in \mathbb{N}$ be given, and assume the lemma holds for all $n' < n$. If $\sum_{m=n}^{\infty} b_m = 0$ then $b_m = 0$ for all $m > n$, which implies $a_m = 0$ for all $m > n$ as well and the lemma holds for $n$. So, we may assume $\sum_{m=n}^{\infty} b_m > 0$. Not further than if $b_n = 0$, then $a_n = b_n \left(\sum_{m=n}^{\infty} a_m\right) / \left(\sum_{m=n}^{\infty} b_m\right) = 0$ and so the lemma holds for $n$ as well. Hence, we may assume $b_n > 0$. Then:

$$\begin{aligned} \frac{a_n}{b_n} &= \frac{\sum_{m=n}^{\infty} a_m}{\sum_{m=n}^{\infty} b_m} \\ &= \frac{\sum_{m=0}^{\infty} a_m - \sum_{m=0}^{n-1} a_m}{\sum_{m=0}^{\infty} b_m - \sum_{m=0}^{n-1} b_m} \\ &= \frac{\lambda \sum_{m=0}^{\infty} b_m - \sum_{m=0}^{n-1} \lambda b_m}{\sum_{m=0}^{\infty} b_m - \sum_{m=0}^{n-1} b_m} \\ &= \lambda \end{aligned}$$

The converse, that if the sequences are proportional then the partial sums are as well, is easy to see.

□

## References


[1] Nick Bostrom. *Anthropic Bias*. Routledge, 2002.
[2] Nick Bostrom and Milan Cirkovic ed. *Global Catastrophic Risks*. Oxford University Press, 2011.
[3] J. R. Gott. Implications of the copernican principle for our future prospects. *Nature*, 363:315–9, 1993.
[4] J. Leslie. Is the end of the world nigh? *American Philosophical Quarterly*, 40:65–72, 1990.
[5] J. Leslie. *The End of the World: The Ethics and Science of Human Extinction*. Routledge, 1996.
[6] Ken Olum. The doomsday argument and number of possible observers. *Philosophical Quarterly*, April 2002.