The Great Filter, Branching Histories and Unlikely Events

David J. Aldous

University of California
Department of Statistics
367 Evans Hall # 3860
Berkeley CA 94720-3860
aldous@stat.berkeley.edu

December 17, 2010

Abstract

The Great Filter refers to a highly speculative theory that implicitly claims one can test, via a probability model, whether known aspects of the history of life on Earth are consistent with the hypothesis that emergence of intelligent life was very unlikely. We describe the theory and some of the many objections to it. We give a mathematical argument to show that one objection, namely that it considers only a single possible linear history rather than multitudinous branching potential pathways to intelligence, has no force.

Key words, Great filter, Drake equation, Fermi paradox, order statistics, rare events, conditioning paradox.

MSC classifications; 60J85, 92D15

1 Introduction

> There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact. *Mark Twain.*

My upper division students enjoy occasional speculative material as a vacation from the usual style of mathematics instruction, so perhaps readers will also. The mathematics used in this article is at the level of exercises in a post-calculus course in mathematical probability.

The generally accepted answer to the question

> have we observed evidence of extraterrestrial intelligence, either contemporary (e.g. via radio astronomy) or in the past (e.g. evidence of previous visits to the solar system)?

is "No". From this single *bit* of data, much speculation has resulted, under the names *Fermi paradox* and *Drake equation.* (The Drake equation is an elaboration of (1) below; that claimed plausible values of its terms lead to large estimates of the number of extraterrestrial civilizations is the Fermi paradox). The topic is discussed at length in two non-technical books [2, 8] and summarized in a well-written Wikipedia article [9], so let me only briefly recall a central style of argument. Consider the product

$$Npq \tag{1}$$

where

- N is the number of Earth-like (loosely, and at formation) planets in the galaxy
- p is the chance that, on such a planet, an intelligent species at a technological level comparable to ours will arise at some time
- q is the chance that such a species would survive in such a way as to be observable (via communication or exploration) to other galactic species for an appreciable length of time.

More precisely, to define q one should weight by "proportion of time" and "proportion of space", relative to lifetime and size of galaxy, that the species would be observable, but such details are not important in our brief account. The point is that Npq represents the number of *other* intelligent species we expect to observe in the galaxy. Because we don't observe any, we conclude *prima facie* (treating absence of evidence as evidence of absence) that it cannot be true that $Npq \gg 1$. Since it would be a bizarre coincidence if $Npq \approx 1$, we should conclude that $Npq \ll 1$ and so humans are most likely to be the only technological species in the galaxy.

The issues in the argument above have been developed in many different directions in the literature. Let me summarize, by slightly caricaturing, three views of the topic. The first two may well be familiar to the reader but the third, the starting point for this article, may not be.

The science fiction view. The bar scene in *Star Wars* is cool; it's nice to imagine there are aliens we could communicate with; so let's think of reasons why the *prima facie* conclusion might be wrong, in other words how there might be extraterrestrial intelligence that we don't observe.

The hard science view. One can hope to estimate N fairly accurately, but there is no conceivable way that either p or q could be estimated scientifically. So further theoretical speculation is futile, though devoting a very small proportion of the science budget to SETI (Search for Extra-Terrestrial Intelligence – an umbrella term for scientific activities seeking relevant data) may be reasonable.

A third view. Human beings did not create the Universe or direct the course of evolution, so N and p are not our responsibility. But q, as applied to us (i.e. will our species leave its mark on the galaxy?) is presumably under our control. Viewing q very roughly as the chance that a hypothetical technological species arising across the galaxy 25 million years in the future would then be able to observe the then-current or previous existence of humans, being told that $q = 10^{-6}$ would be rather depressing. Depressing, because of the ways this might come about, for instance if humans soon become extinct, or change and cease to interact with the macroscopic physical universe. Knowing $q = 10^{-6}$ would be knowing that something like this is *almost certain* to happen. Now having decided that Npq is small, implying pq is very small, the only way to avoid the depressing possibility of q being very small is to for p to be very small.

So this third view leads to the counter-intuitive conclusion that

$$\text{we should } \textit{want } p \text{ to be very small} \tag{2}$$

where the sense of *want* is, as above, "to be consistent with humanity surviving long enough to have at least a tiny chance of leaving its mark on the galaxy".

Given the vast divergence of opinions regarding the value of p which have been expressed by authorities [2, 9], it seems prudent to accept one part of the hard science view:

$$\text{there is no conceivable way that } p \text{ could be estimated scientifically} \quad (3)$$

and have the estimate generally accepted. (Detailed scientific arguments that p is very small are often referred to as the *rare Earth hypothesis* [11]). But note that there is no logical contradiction between believing (3) and being willing to consider the program

> try to ascertain whether what we know about the evolution of life
> on Earth is *consistent with* the possibility that p is very small. (4)

The question is: can we actually carry out anything like this program?

2 The Great Filter

The phrase *Great Filter* was coined in the 1998 Robin Hanson essay [3]. Hanson pointed out the counter-intuitive conclusion (2) and noted the mathematical argument in section 3 as relevant to the program (4). See [1] or [2] pp. 86–90 for other discussions; in this case the Wikipedia article [10] is not the most helpful. There are further counter-intuitive corollaries, for instance [1] that we should *not want* to discover that life ever evolved on Mars independently of life on Earth. The *filter* metaphor is that, in a sequence of steps from non-life to a species detectable from outside its solar system, either the occurrence of some one particular step, or of a sequence of steps in the necessary order, must be incredibly unlikely, otherwise the galaxy would be teeming with observable life.

Let me set up the key mathematical idea via a story with more familiar ingredients. Consider two unlikely events for an individual – say, being struck by lightning and winning the lottery. Suppose I point to a 50-year-old man and tell you that sometime since age 20 he won the lottery and sometime later he was struck by lightning. And suppose the relative probabilities of these two events are not known, except that each is very unlikely. What can you say about the probability distribution of the ages at which these two events occurred?

In thinking about this question with (non-mathematical) intuition, you might think as follows. The two events split the interval $[20, 50]$ into three sub-intervals. In general you have to wait longer for a more-unlikely event than for a less-unlikely event, so the interval ending with the less-likely of the two events will probably be longer.

But, within the simplest probability model one might devise for this story, that conclusion is wrong. Counter-intuitively, *regardless of the relative probabilities*, the three intervals have the same mean length. See section 3.

Hanson [3] pointed out that one can try to apply the same argument to the evolution of intelligent technological species on Earth. Suppose we identify some key steps between the formation of an Earth-like planet and technological civilization. He suggested six intermediate steps (I quote five of his and have modified the sixth).

- Reproductive something (e.g. RNA)
- Simple (prokaryotic) single-cell life
- Complex (archaeatic & eukaryotic) single-cell life
- Sexual reproduction
- Multi-cell life
- Animal-level intelligence
- Technological civilization

The relevant time interval is the interval when life on Earth is possible, roughly 4 billion years ago to 1 billion years in the future. If we suppose each of these 7 steps was an unlikely random event then, under the simplest probability model, the section 3 argument shows that *regardless of the seven actual probabilities*, the eight sub-intervals would have random lengths with equal means. And the data we have on the dates of these events is quite consistent [2] with this model prediction, and so in particular is quite *consistent with* the possibility that p is very small.

It is easy to think of numerous objections to the argument above, and in class I ask students to suggest some. One might object, on philosophical or religious grounds, to any notion that the existence of the human species is the result of randomness. One might object that we have no scientific understanding of how the first few steps happened, so *a priori* no model could be reasonable. One might object to the choice of biological "key steps". One might object to the particular probability model, and so on.

This article will not consider most of these objections, as they are what makes the theory highly speculative, and are not susceptible to mathematical analysis. Instead we focus on another objection, which is that the argument presumes the stated steps are "the single path to intelligence". It is sometimes argued [5] that *convergent evolution* implies that there were only relatively few available pathways for evolution on Earth, but this is a minority viewpoint amongst evolutionary biologists. And especially when considering hypothetical independent evolution of life in other solar systems under somewhat different physical conditions, my personal intuition suggests that there might well be entirely different pathways whose nature we cannot even imagine.

Section 4 sets up and analyzes a model allowing an arbitrarily large number of possible branching paths to technological intelligence. Regardless of the number and structure of possible pathways and regardless of the chances for each step on each path to actually occur, is it still true that, conditional on the (assumed unlikely) event that some path does lead to technological intelligence, the times at which the key steps may occur are still uniform random? We show that the answer is "yes, essentially". Section 5 discusses two other variant models, and section 6 provides a final reminder of the problematic nature of this modeling endeavor.

3 The single path model

We first recall some elementary material. Take n independent random variables with uniform distribution on $[0, 1]$ and then either re-order them into increasing order, or condition on their being in increasing order. The resulting random variables $0 < X_1 < \ldots < X_n < 1$ have joint (probability) density function

$$f_X(x_1, \ldots, x_n) = 1/n!, \quad 0 < x_1 < \ldots < x_n < 1, \tag{5}$$

that is to say are uniform on the space of possible values. Moreover each of the $n + 1$ random variables giving interval lengths

$$X_1,\ X_2 - X_1, \ldots, X_n - X_{n-1}, 1 - X_n$$

has the same distribution, with density

$$g(u) = n(1 - u)^{n-1}, \quad 0 < u < 1$$

and mean $1/(n + 1)$.

The Exponential(λ) distribution has density

$$f(u) = \lambda e^{-\lambda u},\ 0 < u < \infty$$

and models the time of first occurrence of an event which is ‘purely random” in the sense of having chance λdt of occurring in a time interval of length dt, independent of previous events (in less precise language, an event that occurs at rate λ). Given $\lambda_1, \ldots, \lambda_n$, consider random times $0 < Z_1 < \ldots < Z_n < \infty$ such that the random variables $Z_1, Z_2 - Z_1, \ldots, Z_n - Z_{n-1}$ are independent and $Z_i - Z_{i-1}$ has an Exponential(λ_i) distribution. This is the “prior” or “unconditional” version of the single path model; with n specified steps such as those in section 2, and with unlimited time, the i'th step would occur at time Z_i. So we are assuming that, after the $(i-1)$'th step, the i'th step occurs at rate λ_i. The joint density function of $(Z_1, \ldots, Z_n)$ is then

$$\prod_{i=1}^{n} \lambda_i \exp(-\lambda_i(z_i - z_{i-1})), \quad 0 < z_1 < z_2 < \ldots < z_n < \infty.$$

Now condition on the event (which will be unlikely if all the λ_i are small) that all steps occur before time 1, in other words that $Z_n < 1$. We choose the time unit to be the length of interval of interest, and condition on all steps being completed before this time. Writing $0 < Y_1 < \ldots < Y_n < 1$ for the occurrence times of the steps under this conditioning, their joint density is

$$f_Y(y_1, \ldots, y_n) = c \exp(-\sum_{i=1}^{n} \lambda_i(y_i - y_{i-1})), \quad 0 = y_0 < y_1 < \ldots < y_n < 1 \tag{6}$$

where $c = \frac{\prod_i \lambda_i}{P(Z_n < 1)}$. This is the “posterior” or “conditional” version of the single path model, for the times at which steps occurred, given that they all occurred before time 1.

In the Great Filter argument we are assuming each λ_i is small. We can now argue qualitatively as follows. The exponential term in (6) is ≈ 1, that is

$$f_Y(y_1, \ldots, y_n) \approx c, \quad 0 = y_0 < y_1 < \ldots < y_n < 1 \tag{7}$$

which implies that the distribution of $(Y_1, \ldots, Y_n)$ is approximately the uniform distribution (5). So the conclusion is that

$$\text{the times } Y_1,\ Y_2 - Y_1, \ldots, Y_n - Y_{n-1}, 1 - Y_n \text{ between steps} \\ \text{all have approximately the same distribution,} \tag{8}$$

with mean approximately $1/(n+1)$. This holds regardless of the actual values of the λ_i, other than being assumed small, and is the mathematical content of Hanson's [3] argument.

We will refer to the quantities $Y_1,\ Y_2 - Y_1, \ldots, Y_n - Y_{n-1}, 1 - Y_n$ as *inter-step times.*

3.1 Comments

The normalizing constant c in (6) must approximate the normalizing constant $1/n!$ in (5), so we obtain the approximation

$$P(Z_n \leq 1) \approx \frac{\prod_i \lambda_i}{n!}$$

for the unconditional probability p in our model. This of course *does* depend very much on the actual values of the λ_i. Like many of the well-known counter-intuitive results in mathematical probability (e.g. inference from a positive diagnostic test result for a rare disease), our intuition about the effect of conditioning can be poor.

There is no great difficulty in replacing these slightly vague approximations by rigorous inequalities involving $\max_i \lambda_i$, but we leave this to the reader with a taste for calculus approximations. Instead we want to consider the issue of whether the counter-intuitive result (8) holds only in this precise setup or whether it is robust under certain variant assumptions. As already mentioned, the variant on which we focus is replacing the "single possible path" by "multiple branching possible paths". This issue of course arises in general discussion of the Fermi paradox, but apparently has not been discussed in the specific mathematical context of the Great Filter argument.

4 The branching history model

Here the mathematical setup is a tree, where the root (the top vertex in Figure 1) represents the original "lifeless newly-formed Earth-like planet", and each "child" vertex v represents a different possible "first step" on a path of steps. For each such v there are different possible "second steps" w, and so on, and each downward path leads (after some path-dependent number of steps) to a "leaf" representing a technological species. For each pair (v, w) of possible successive steps there is a constant λ_w, and if step v occurs, then the rate at which step w subsequently occurs equals λ_w. For each v, when the first of these possible next steps occurs, the other possibilities are suppressed. So in unlimited time, there would eventually be one random path, with a random number L of random steps from the root $= v_0$, through $V_1, V_2, \ldots$ to a random leaf V_L, the successive steps happening at random times $0 = T_0, T_1, \ldots, T_L$. This is the "prior" or "unconditional" version of the branching history model.

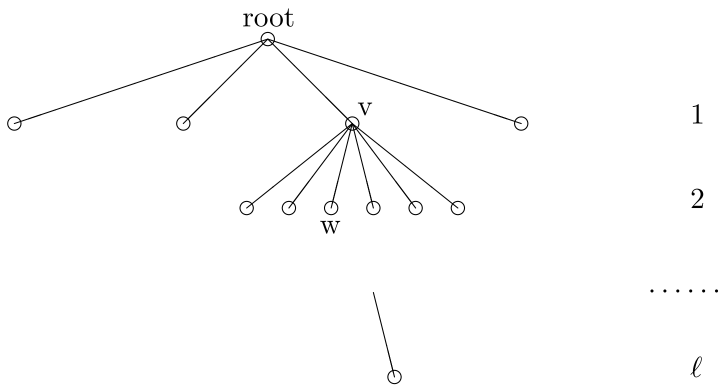

Figure 1. Branching possible histories.

We can now state the question. Assuming the chance p that this random path reaches a leaf before time 1 is small, but conditioning on this happening, and thus observing the value of the number of steps L, it it true (as in the single path model) that the $L+1$ inter-step times have approximately the same distributions, regardless of the shape of the tree and regardless of the actual values of the λ_v (assumed small)?

Note we are implicitly assuming that the number of steps to a leaf should typically be small (e.g. 7, as in the section 2 list) but the branching numbers, namely the numbers of alternative next steps after each step, can be arbitrarily large.

What does intuition suggest? On the one hand, one might argue that because there can only be one actual realized path, the fact that this path is random rather than prespecified is hardly important, so that this model is not essentially different from the single path model. On the other hand one can think of possibly analogous settings where such arguments are incorrect. Recall a freshman statistics example. If, hypothetically, 50% of families had one child and 50% had three children, then the average number of children equals 2, from the parents' viewpoint. But 75% of children are in three-child families, so the average number of children equals 2.5 from the children's viewpoint. In looking at the genealogy of a random living person, this effect continues backwards indefinitely; the 10'th generation great great ... great grandfather cannot be a statistically typical member of his generation, because we know he had a son for sure. So by analogy, the times of intermediate steps along the random path might not be statistically typical.

Anyway, a calculation is required to answer the question.

4.1 Analysis of the branching history model

We first quote further elementary facts about the Exponential(λ) distribution. Given m independent events where the i'th event occurs at rate μ_i, that is m independent random times with Exponential(μ_i) distributions, the random time T and random index I of the first event to occur are such that

$$P(T \in [t, t+dt], I = i) = \mu_i \exp(-\mu^* t)\ dt; \qquad \mu^* := \sum_j \mu_j \tag{9}$$

which is equivalent to the three properties

$$T \text{ has Exponential}(\mu^*) \text{ distribution}$$

$$P(I = i) = \mu_i/\mu^*, 1 \le i \le m$$

$$T \text{ and } I \text{ are independent.}$$

We can now readily write down the probability distribution in the unconditional branching history model, by simply applying (9) conditionally at each step. For any path to a leaf starting with the root v_0, say $v_0, v_1, v_2, \ldots, v_\ell$, and any $0 = t_0 < t_1 < \ldots < t_\ell < \infty$,

$$P(L = \ell; V_i = v_i, 1 \le i \le \ell; T_i \in [t_i, t_i + dt_i]) = \prod_{i=1}^{\ell} \lambda_{v_i} \exp(-\lambda^*_{v_{i-1}}(t_i - t_{i-1}))\ dt_i \tag{10}$$

where

$$\lambda^*_v := \sum \{\lambda_w : \ w \text{ a possible next step after } v\}.$$

We now see that to repeat the argument from the single path model, what we need to assume is that

$$\text{each } \lambda^*_v \text{ is small,} \tag{11}$$

rather than just assuming each λ_v is small. So we are assuming that after each step, the chance that *some* next step occurs is small.

Recall that T_L is the time at which some leaf is reached. For the conditioned process we condition on the event $T_L < 1$. Adding tildes to denote conditioned variables, assumption (11) implies that the exponential terms in (10) are ≈ 1. So for the conditioned process we have

$$P(\tilde{L} = \ell; \tilde{V}_i = v_i, 1 \le i \le \ell; \tilde{T}_i \in [t_i, t_i + dt_i]) \approx c \prod_{i=1}^{\ell} \lambda^*_{v_{i-1}}\, dt_i; \quad 0 < t_1 < \ldots < t_\ell < 1.$$

The normalizing constant is $c := 1/P(T_L < 1)$. Note that this constant, and $c_2(\ell), c_3(\ell)$ below, depend on the tree and the whole configuration of λ-values on the tree, but not on the particular values of v_i or t_i on the left hand side. Summing over all paths of a given length ℓ gives

$$P(\tilde{L} = \ell; \tilde{T}_i \in [t_i, t_i + dt_i]) \approx c_2(\ell) \prod_{i=1}^{\ell} dt_i; \quad 0 < t_1 < \ldots < t_\ell < 1,$$

while conditioning on the value of $\tilde{L}$ gives

$$P(\tilde{T}_i \in [t_i, t_i + dt_i] \mid \tilde{L} = \ell) \approx c_3(\ell) \prod_{i=1}^{\ell} dt_i; \quad 0 < t_1 < \ldots < t_\ell < 1.$$

This is the analog of (7) in the single path model, and implies that conditional on $\tilde{L} = \ell$ the $\ell + 1$ inter-step times $\tilde{T}_{i+1} - \tilde{T}_i$ have the same mean, $1/(\ell + 1)$.

5 Other variants of the model

5.1 Time-dependent rates

The model above assumed not only that, given previous steps, the waiting time until the occurrence of a next step was random, but also that the rate (probability per unit time) did not vary with the time t since the previous step. Suppose, in the single path model (the branching case is similar), that for the i'th step we replace the constant rate λ_i by a time-dependent rate $\mu_i(t)$. Loosely speaking, imagine that background conditions change deterministically so as to make occurrence of the (still unlikely and random) next step become *relatively* more likely as time t since previous step increases. The mathematically simplest increasing function is the linear function, so take

$$\mu_i(t) = \lambda_i t \tag{12}$$

for small constants λ_i.

It is easy to repeat the analysis where each step i is assumed to occur at a rate of form (12) for constants λ_i. It turns out that the basic conclusion, as stated in (8), remains true: conditioned mean inter-step times are approximately equal (the joint distribution changes from (7) but it retains the desired symmetry property, technically called *exchangeability*).

The argument really depends only on the rates $\mu_i(t)$ for each step having the same functional form (linear, in this case) with differing multiplicative constants. But if the functional form varies between steps then we would lose symmetry, and conclusion (8) would break down.

5.2 Climbing a gentle staircase of complexity

Any way of addressing the overall program (4) must contain some probability model for the appearance of technological species. But arguably the whole "Great Filter" scenario reflects a common perceptual bias toward thinking of chance in terms of unlikely events with dramatic consequences (recall the iconic examples of winning the lottery or being struck by lightning), a bias perhaps reinforced by the recent best-seller *The Black Swan* [7]. An opposite perception is to think of chance entering via the cumulative effect of many, individually minor and not unlikely, chance events. Imagine a casino gambler starting with 50 dollars, betting 1 dollar on red or on black at roulette, and staying until his fortune reaches 0 or 500 dollars. In the (unlikely) event of reaching 500 dollars, a textbook exercise shows that (up to endpoint corrections) the sequence of wins and losses follows almost precisely (that is, up to endpoint corrections) a process in which the chance of winning is 20/38, rather than the true chance 18/38. So imagine evolutionary history as a long but not steep staircase; from one level of complexity (of individual creatures or of ecosystems) the system might next go up to a slightly higher level or down to a slightly lower level. Conditionally on reaching a high level (technological civilization), what we observe retrospectively is an overall increase; but it might be that the natural tendency is always for decrease to be more likely than increase, and that what we observe is the result of a very unlikely overall process, analogous to our lucky roulette gambler.

In many ways this is a much more disconcerting scenario, namely that the whole observed pattern of evolutionary history might be systematically biased away from what was *a priori* likely. This does not seem to be statistically testable.

6 Beware of simplistic models

It is surely a truism that, in the vast majority of real-world contexts where we perceive chance as playing a role, what is actually happening is a complex mixture of random and deterministic effects that is not amenable to simplistic mathematical modeling at all. So one needs to be very skeptical of work like ours! Indeed we can demonstrate why, by thinking a little less superficially about our example of winning the lottery and being struck by lightning. Our model assumes that chances, while varying between individuals, are constant in time (e.g. per year) for an individual. Is this true? There is extremely accurate and accessible data [6] on *deaths* by lightning in the U.S., showing that young adult males are relatively more at risk. Presumably this holds also for being *struck* by lightning, so the "constant rate in time" assumption is far from true. Similarly, it is known [4] that big lottery winners in the U.S. are on average older than the general adult population, so again the "constant rate in time" assumption is hardly true. Hence the need for healthy skepticism.

Acknowledgment. I thank an anonymous referee for numerous expository improvements.

References

[1] N. Bostrom. Where are they? Why I hope the search for extraterrestrial life finds nothing. *Technology Review*, pages 72–78, 2008. http://www.nickbostrom.com/extraterrestrial.pdf.

[2] P. Davies. *The Eerie Silence: Renewing Our Search for Alien Intelligence*. Houghton Mifflin Harcourt, Boston MA, 2010.

[3] R. Hanson. The great filter - are we almost past it?, 1998. http://www.webcitation.org/5n7VYJBUd.

[4] H.R. Kaplan. Lottery winners: The myth and reality. *J. Gambling Studies*, 3:168–178, 1987.

[5] S. Conway Morris, editor. *The Deep Structure of Biology: Is Convergence Sufficiently Ubiquitous to Give a Directional Signal?* Templeton Press, West Conshohocken, PA, 2008.

[6] National Weather Service. Lightning safety statistics, 2010. http://www.lightningsafety.noaa.gov/more.htm.

[7] N. N. Taleb. *The Black Swan: The impact of the highly improbable*. Random House, 2007.

[8] S. Webb. *If the Universe Is Teeming with Aliens . . . Where Is Everybody?* Springer, New York NY, 2002.

[9] Wikipedia. Fermi paradox — wikipedia, the free encyclopedia, 2010. [Online; accessed 16-June-2010].

[10] Wikipedia. Great filter — wikipedia, the free encyclopedia, 2010. [Online; accessed 17-June-2010].

[11] Wikipedia. Rare earth hypothesis — wikipedia, the free encyclopedia, 2010. [Online; accessed 30-June-2010].