Bayesian Estimate of an Infinite/Infinitesimal Universe

The Infinite:

A Bayesian Estimate

How can we apply Bayes' theorem to determine the probability of the Universe being infinite? First step is to establish the priors. We'll call the hypothesis p(I), the probability of an infinite Universe and the alternative hypothesis p(~I), the probability of a finite universe, which rationally covers all alternative possibilities. Now we need some evidence that we think supports the idea of an infinite Universe. What is it that makes it seem endless? What can we point to as evidence? Well, the Universe is certainly very big, but big compared to what? It's very big compared to what we used to think a long time ago. And it's been bigger every time we ever looked hard enough to find a more distant object. So our evidence could be stated as, D, the persistent discovery of an ever more distant object.

Under the above criteria we can start with the history's first discovery of a very distant object and calculate Bayes Theorem the first time using our best (subjective) guess for the priors. Then, every time we discover another, more distant object that implies a larger Universe we can use the previously calculated *posterior probability* as rational basis for the *prior probabilities* in our next calculation.

* * *

We start out with our first unknown posterior probability, p (I|D), as the probability of the Universe being infinite, I, (the phenomenon in question) given discovery, D, of a more distant object (the evidence at hand).

$$p(I|D) = \frac{p(D|I)*p(I)}{p(D|I)*p(I) + p(D|{\sim}I)*p({\sim}I)}$$

Thus, for our priors we have assignments:

p (D|I) = The probability of discovering a more distant object D, assuming that the Universe is infinite (*inverse conditional probability*). This probability, by definition, would be 1 or true. If the Universe is infinite we will necessarily *always* be able to find some more distant object. As we shall see, it is the unarguable assignment of unity to this variable which helps anchor the equation's behavior.

p (I) = The prior probability of the Universe being infinite *independent* of finding a more distant object (*the prior probability of* I). If we started with Aristotle he would have probably assigned a near-zero probability to this variable since we know that he thought infinity was only a conceptual quantity. But, on the other hand, if Aristotle were sitting here today, and found out how big the Universe is so far, with no end in sight, he might lean more towards assigning near-certainty to the prior probability of infinity. So, for now, we will take the liberty of having Aristotle go along with a 0.5 probability (50/50 split between p(I) and p(~I)) that Bayesians generally assign to priors for very subjective, uncertain, competing phenomena. This way the evidence alone will be responsible for favoring one over the other.

p (~P) = The prior probability of the Universe being *finite* (*normalizing alternative probability*). This gets the other half of the 50/50 split, or 0.5.

p (D|~I) = The probability of discovering a more distant object D, given that the Universe is *finite* (the *normalizing conditional probability*). We will subjectively assign that value a near-true probability, say, 0.999999, to allow for the fact that the universe must be somewhat bigger than we already know and thus we are very likely to find a more distant object before running out of Universe, even if it is finite. However, keep in mind, that the larger we make this probability assignment for fairness sake, the larger we are implying that the universe will turn out to be. At 0.999999 we are basically saying that the Universe will always be a million times larger than whatever we can see of it right up until we bump into the end of it. We can't go all the way to 1 for fairness sake, because that would be the same as saying that the Universe is infinite. Furthermore, we will not decrease this value every time we discover more distant object, even if we find one that is a million times further away than our starting point. This is an extremely generous allowance for the possibility of a *finite* universe. The fact that we have to assign some value short of 1 to p(D|~I) is the other factor in this particular arrangement of Bayes' theorem that constrains its behavior. The 50/50 split between D and ~D turns out to be ultimately immaterial to the outcome.

So we calculate our first iteration on the steps of the Parthenon with Aristotle looking over our shoulder.

$$p(I|D)_0 = \frac{(1)(0.5)}{(1)(0.5) + (0.999999)(0.5)}$$

$$P\,(I|D)_0 = 0.50000025$$

Thus, the initial calculation of the probability of the Universe being infinite is 50.000025%, which is just a split frog-hair better than the 50/50 arbitrary selection of the priors. Aristotle is unimpressed. Not much of a ringing endorsement for an infinite Universe, he says. Well, actually, it is. This is just the first of 2 trillion iterations.

In Aristotle's time, the general consensus is that the stars are all hung on the outermost crystal sphere beyond the Sun and all the planets. Now Aristotle didn't even know how far away the moon was nor perhaps even what a billion kilometers might amount to, but we will grant him an extremely generous initial size for his Crystal Sphere Universe of 16 billion kilometers (or 10 billion miles) a span which easily accommodates the orbits of all eight planets and pluto. As outlined above, we will then run history forward, and for every time some distant object's discovery adds another 16 billion kilometers, we will recalculate Bayes Theorem with the *new* prior probability that the Universe is infinite, p(I), being assigned from previously calculated value of the Universe being infinite, p(I|D). In this way we update our Bayesian estimate with each new discovery lending a little more weight to our prior assessment of p(I).

Also, from the fact that p(I) and p(~I) are mutually exclusive propositions, their combined probability must necessarily equal 1 (initially 0.5 and 0.5). Thus, if the new prior assignment of p(I) from the initial calculation is to be 0.50000025, p(~I) must be 1-0.50000025 or 0.49999975 for the next iteration. Again, erring on the conservative side, we will leave p(D|~I) (the likelihood of discovering a more distant object, presuming the Universe is *finite*) unperturbed at 0.999999 instead of reducing its likelihood a little bit every time another 16 billion km is added to the known size of the Universe.

This iterative process is known in mathematics as a sequence. And it is a powerful tool by which to integrate the incremental updating behavior of Bayes Theorem over a long series of ongoing evidence.

Thus, the Bayesian equation,

$$p(I|D) = \frac{p(D|I)*p(I)}{p(D|I)*p(I) + p(D|\sim I)*p(\sim I)}$$

becomes the Bayesian sequence,

$$p(I|D)_n = \frac{p(D|I)*p(I|D)_{n-1}}{p(D|I)*p(I|D)_{n-1} + p(D|\sim I)*(1-p(I|D)_{n-1})}$$

where "n" denotes how many times we have incremented the sequence each time we add another 16 billion kilometers to the size of the known universe.

By the mid 1800s, a Scott by the name of Thomas Henderson using the parallax (triangulation) displacement against more distant background stars offered by the width of Earth's orbit around the Sun, discovered that the nearest visible star, Alpha Centauri, is about 4.3 light years (40,678,000,000,000 km to Thomas' way of thinking) distant, and we're off to the races. Since this eclipses the 16 billion km crystal sphere by a factor of 2400 by the new radial distance to Alpha Centauri, we can plug the new priors and cycle our sequence as follows:

$$p(I|D)_0 = 0.50000025$$

$$P(I|D)_1 = \frac{p(D|I)*p(I|D)_0}{p(D|I)*p(I|D)_0 + p(D|\sim I)*(1-p(I|D)_0)}$$

$$p(I|D)_1 = \frac{(1)(0.50000025)}{(1)(0.50000025) +(0.999999)(0.49999975)}$$

$$p(I|D)_1 = 0.5000005$$

$$p(I|D)_2 = \frac{p(D|I)*p(I|D)_1}{p(D|I)*p(I|D)_1 + p(D|\sim I)*(1-p(I|D)_1)}$$

$$p(I|D)_2 = \frac{(1)(0.5000005)}{(1)(0.5000005) + (0.999999)(1-0.5000005)}$$

$$p(I|D)_2 = 0.50000075$$

And so on for 2400 iterations whereupon we arrived at:

$$p(I|D)_{2400} = 0.50060025$$

Over the course of 2400 iterations we come up with a slightly higher (a little over 0.1% overall increase) probability of the universe being infinite. Not a ringing endorsement for an infinite universe. But that's just the first of billions of iterations.

Turns out, our sequence approaches certainty long before we iterate through even a tiniest fraction of the *visible* universe. Even before we exit our own insignificant galaxy's diameter of 100,000 light years, the equation has already converged to virtual certainty for an infinite Universe. And we've only covered 1/100,000 of the part of the universe we can see from Earth.

$$p\ (I|D)_{59\ \text{million}} \cong 1$$

In fact, even going back and reassigning the priors with more radically negative initial subjective opinions with respect to an infinite Universe doesn't change the outcome. Giving a value of a billion to one against the probability of an infinite Universe (p (I) = 0.000000001), and a billion to one in favor of the probability of a *finite* Universe (p (~I) = 0.999999999), and an even more finite-biased billion to one chance, given a *finite* Universe, of always finding a more distant object, and not decreasing it ever, even after we've found a billion more distant increments, (p(D|~I) = 0.999999999), the equation still converges to 1 or true for the probability of an infinite universe.

Somewhere well beyond the edge of our Galaxy (but nowhere near our nearest galactic neighbor) the convergence of the sequence begins accelerate dramatically.

$$pp\ (I|D)_{19\ \text{billion}} = 0.1514$$

$$pp\ (I|D)_{20\ \text{billion}} = 0.3266$$

$$pp\ (I|D)_{25\ \text{billion}} = 0.9863$$

Before we get a $1/20^{th}$ of the way to Andromeda, our partner galaxy and closest intergalactic neighbor, just 2.2 million light-years away, the sequence has converged to true, for an infinite Universe: .

$$pp\ (I|D)_{30\ \text{billion}} = 0.999906433$$

$$pp\ (I|D)_{40\ \text{billion}} = 0.999999996$$

$$pp\ (I|D)_{50\ \text{billion}} \cong 1$$

Not only can we see Andromeda with the naked eye, it is closer to us than the edge of the visible Universe by a factor of over a million leaving around 2 trillion iterations on our equation left to go. If you would like to play with the equation using a calculator, it turns out that the sequence is a special case of the Riccati recursive relation and the following is a more accessible solution for a given number of iterations, n:

$$p(I|D)_n = \frac{p(I|D)_0}{p(I|D)_0 + (1-p(I|D)_0)(p(D|\sim I)^n)}$$

* * *

Given that $p(D|I) = 1$, our Bayesian sequence has the inherent property of being convergent to 1 (see last section of this paper) for the existence of an infinite Universe, no matter what we do with the subjectively assignable priors. Whether we iterate every time the Universe grows by 16 billion kilometers, 16 billion light years or even every time it doubles in size, the equation will ultimately converge to true for an infinite Universe.

* * *

The Infinitesimal:

<u>A Bayesian Estimate</u>

So what is the Bayesian perspective on the infinitesimal? For simplicity's sake we can arrange it to be identical in structure to the vast perspective. The same process that suggests an infinitely large Universe is similarly at work in the smaller dimensions. That is, each time we look hard enough for a smaller object, we find one.

The equation sets up with "i" representing the Infinitesimal (instead of the infinite), "D" representing the Discovery of a smaller particle (rather than a more distant object) and p(i|D) representing the probability of the universe being infinity divisible given the discovery of a smaller unit of matter. This arrangement delivers a Bayesian equation for the probability of an infinitely divisible Universe identical to that of our infinitely large analysis:

$$p(i|D) = \frac{p(D|i)*p(i)}{p(D|i)*p(i) + p(D|\sim i)*p(\sim i)}$$

And its Bayesian sequence:

$$p(i|D)_n = \frac{p(D|i)*p(i|D)_{n-1}}{p(D|i)*p(i|D)_{n-1} + p(D|\sim i)*(1-p(i|D)_{n-1})}$$

Thus, every time we discover a smaller particle we iterate through the number of times that particle can be divided into our 1 micron sized ceramic dust particle starting point. The key conditional priors of p(D|i) and p(D|~i) are assigned with the same values from our vast analysis of 1 and 0.999999 respectively while p(i) and p(~i) spilt the initial prior probability field with 0.5 each. Thus, we have the same overall behavior of both the equation and the sequence as we had with the infinite case, with the final results inevitably converging to 1 or true in favor of an infinitely divisible Universe given sufficient iterations.

Our starting point for a baseline unit of tinyness might be on the order of the grain size of the finest ceramic powders of Aristotle's day of 1 micron (one millionth of a meter). Starting our calculations with the first iteration based on discovering something smaller than our initial one micron (10^{-6} m) ceramic powder grain size from the pottery mills of Aristotle's day we have:

$$P(i|D)_0 = \frac{(1)(0.5)}{(1)(0.5) + (0.999999)(0.5)}$$

$$p(i|D)_0 = 0.50000025$$

Then using our Riccati recursive solution for the sequence, we have:

$$P(I|D)_n = \frac{0.50000025}{0.50000025 + (1-0.50000025)(0.50000025)^n)}$$

Solving sequence down through the molecular and atomic scales to the size of the smallest atom, the hydrogen atom at 10^{-10} m, which is 10,000 times smaller than our one micron starting point, for n = 10,000 iterations of our sequence. At this point our probability calculation is:

$$p(i|D)_{10,000} = 0.5025$$

Not much of an improvement over our 0.5 prior, but the atom turns out to be pretty big compared to the nucleus. Running it down through a hundred million iterations to 10^{-14} m size of the hydrogen nucleus, we converge to 1 or true for an infinitely divisible Universe:

$$p(i|D)_{100\ \text{million}} \cong 1$$

So, we discover that the probability of an infinitely divisible Universe is true well before we iterate down to the size of a proton. Yet, we are still 999.9 billion iterations short of the upper limit of the size of an electron (10^{-18} m). So, for fairness sake, we start over and reassign the priors with the extremely finite-biased values.

We will start over with a value of a one in a billion against the probability of an infinitely divisible Universe (p (i) = 0.000000001), and a billion to one in favor of the probability of a *finite* Universe (p (~i) = 0.999999999), and a billion to one chance of always finding a smaller quantum object even in a finite universe, and not decrementing it at all, even if we find a billion smaller objects, (p(D|~i) = 0.999999999).

$$pp\ (i|D) = \text{pessimistic probability}$$

$$pp(i|D)_0 = \frac{(1)(0.0000000001)}{(1)(0.0000000001) + (0.999999999)(0.999999999)}$$

Iterating down through the size of the atom:

$$pp\ (i|D)_{10,000} = 0.00000000100001$$

Not much help. Iterating on down through the size of the nucleus:

$$pp\ (i|D)_{100\ \text{million}} = 0.000000001105$$

Little better, but still looking finite. Iterating down to the size of a proton:

$$pp\ (i|D)_{1\ billion} = 0.00000000272$$

Not very promising. But somewhere between the 1/10th and 1/30th the size of a proton the sequence turns an asymptotic corner:

$$pp\ (i|D)_{10\ billion} = 0.000022$$

$$pp\ (i|D)_{20\ billion} = 0.327$$

$$pp\ (i|D)_{30\ billion} = 0.999906$$

$$pp\ (i|D)_{50\ billion} \cong 1$$

* * *

At around 1/50th the size of a proton the sequence converges very nearly to 1 or true, and we are still have 950 billion iterations left out of our original trillion.

As in our vast analysis, no matter how unfavorably we assign the priors, our sequence for the probability of the universe being infinitely small converges to true (see next section) given enough iterations. Now, it's also true that we can keep assigning ever more negative priors that would inhibit the sequence from converging for ever smaller dimensions we might discover, but that would be tantamount to saying that the world gets smaller than we can ever determine, no matter how long investigate, but then stops getting smaller somewhere past that. In light of our ongoing discovery process of ever smaller features of matter, we should be more reasonably inclined to assign priors somewhat more unbiased against the prospect of an infinite continuum.

Convergence

And interesting property that emerges from this particular arrangement of Bayes' Theorem is that of *convergence*. Since in both analyses p(D|I), p(D|i) = 1 and p(D|~I), p(D|~i) < 1, and since all other variables are, as probabilities, confined between the values of 0 and 1, this sequential arrangement of Bayes' Theorem has the inherent property of being *convergent* to true, or $p(I,i|D)_n = 1$, as n approaches infinity no matter how pessimistically one might assign the priors.

Substituting p(I|D), p(i|D) = X, p(D|I), p(D|i) = 1, and p(D|~I), p(D|~i) = α we have,

$$X_n = \frac{X_{n-1}}{\alpha + (1-\alpha)\, X_{n-1}}$$

And

$$X_{n+1} = \frac{X_n}{\alpha + (1-\alpha) X_n}$$

Which is a special case of the Riccati recursive relation:

$$X_{n+1} = \frac{a + bX_n}{c + dX_n}$$

where a = 0, b = 1, c = α and d = 1 – α., which solves in closed form through a bit of acrobatics described at

http://www.hypatia.math.uri.edu/~kulenm/diffeqaturi/chad442/project.htm

from which derives the solution.

$$X_n = \frac{\beta}{\beta + (1 - \beta)\, \alpha^n}$$

Where $\beta = X_0$ or the initial calculation of the probability in question [$p(D|I)_0$, $p(D|i)_0$]. This solution is convergent to 1 for all α and β smaller than 1 and greater than 0. Thus, whatever values are assigned the priors (within the allowed limits) the probability will converge to true for a large enough data set under this arrangement of Bayes' Theorem.

Any corrections or improvements will be greatly appreciated.

Send all inquiries to: author@thegodofreason.com