Bayesian Probability Assessment of the Rules of Discovery

There is a certain recurrent psychological pathology we humans succumb to in our examinations of the cosmos at the largest scale we can see at any give time. We don't realize we are doing it, but every time we expand our cosmological view, we inadvertently presume it could be for the last time. Every time we find some immense structure, we subconsciously imbue that structure with whatever properties necessary to make it a candidate for a complete description of the universe. But every time we think we are observing a final configuration of the universe we eventually find a larger structure under which the full extent of our former configuration becomes a constituent element. The full range of the Earth was found to be but a tiny part of the Solar System; the Solar System, a tiny part of the Milky Way; the Milky Way, a tiny part of the Big Bang.

Concurrent to the largest structure misconception is our inadvertent tendency to characterize a structure as unique. If a structure is seen as a final configuration of the universe, it is also understandably presumed to be unique in the universe. In every case, however, the largest structure we could observe at any given time was determined *not* to be the largest structure by merit of the discovery of other examples of that same class of object scattered across the distant cosmos. The Earth proved to be one of many planets; the Sun, one of many stars; the Milky Way, one of many galaxies. In mitigation of such misplaced faith in the sufficiency any human cosmology, there are two axioms that serve to formalize a more probable Bayesian outcome to our investigations.

The two proposed rules of discovery are stated as follows.

The Finite Rule:

1) *All material phenomena are finite constituents of larger structures.*

The Plurality Principal:

2) *All material phenomena are multiply manifest.*

Both rules were derived by applying Bayes' Theorem to both the broad-scale physical structure of the universe and to the history of the human characterization of the universe. This supplemental paper briefly describes Bayes' Theorem and its specific application for assessing the probability of the above two rules of discovery being true beyond the scale of the Big Bang.

Bayes' Theorem

The root form Bayes' Theorem is stated as follows:

$$p(H|E) = \frac{p(E|H)*p(H)}{p(E|H)*p(H) + p(E|{\sim}H)*p({\sim}H)}$$

In its root form Bayes' theorem is deceptively simple. In application, however, the variables can often be complicated functions integrated across a wide range of incidental data. In this analysis, the variables are held to a minimum complexity for clarity's sake. The p( ) operator indicates the probability of the hypothesis or phenomenon (H, ~H) or the relationships (H|E, E|H, E|~H) inside the parentheses being true. The relationship delineated by the vertical bar "|" is read as the probability the variable to the left of the bar being true *given* or *presuming* that the variable to the right of the bar *is* true. On the left hand side of the full equation, the term p (H|E) stands for the probability of hypthesis H being true *given* the existence of evidence E. This term is what we're looking for: the results of the equation or the *posterior probability*.

Calculations

Using Bayes' Theorem we will examine the probability of the two rules of discovery being true at the level of the Big Bang under the impact of two distinct ranges of data. The first range of data is the structural relationships of matter across the full scalar range of the known universe and the second will be the sequential relationship between successive cosmological theories across the range of human history.

Structural Analysis

Taking the material structure of the universe from quarks to galactic clusters as the evidentiary field we can calculate the probability of both rules of discovery being true for the Big Bang as follows:

$$p(Fm|E) = \frac{p(E|Fm)*p(Fm)}{p(E|Fm)*p(Fm) + p(E|\sim Fm)*p(\sim Fm)}$$

Where p(Fm|E) is the probability of the current material object under considerations being a *finite*, *multiply manifest* constituent of a larger object (Fm), given the evidence that the current object is collectively assembled from smaller objects (E). Thus the priors are assigned their subjective initial values as follows:

p (E|Fm) = The probability of finding evidence E (the existence of smaller objects collectively assembled within the current object) assuming that phenomenon Fm (all objects being finite and multiply constituent to a larger object) is true. This value is necessarily (tautologically) 1 or true.

p (Fm) = The prior probability of hypothesis Fm being true *independent* of the evidence E, or *the prior probability of* Fm. For neutrality's sake we start with a 50/50 split between Fm (all objects are finite and multiply constituent to larger objects) and ~Fm (all objects are *not* finite and multiply constituent) giving Fm = 0.5.

p (E|~Fm) The probability of evidence E (the existence of smaller objects collectively assembled in the current object) being true presuming that our hypotesis Fm (all objects are finite constituents of larger objects) is *not* true. We will assign p (E|~Fm) = 0.5 generously allowing that the hierarchical evidence we see in all matter observed so far has an equal chance of being some how unrelated or contrary to our hypothesis.

p (~Fm) = The prior probability that Fm is *not* true, independent of E, which is necessarily 1 – (Fm) = 0.5 for our initial conditions.

So we take one of the smallest object we can currently detect (the neutron) that we know contains yet smaller objects (quarks), where one class of quark (the up quark) assembles with two quarks of another class (down quarks) to form neutrons and calculate the following Bayesian probability of neutrons being finite constituents of yet larger objects.

$$p(Fm|E) = \frac{1 * 0.5}{(1 * 0.5) + (0.5 * 0.5)}$$

p(Fm|E) = 0.67 (neutron constituent to a larger object)

Which is a modest improvement over the initial subjective assignment of p(Fm) = 0.5. However, as is often the case in Bayesian analyses, we will use the last calculated value of the *posterior* probability as the new *prior* probability for consideration of new evidence. Thus, $p(Fm)_n = p(Fm|E)_{n-1}$ in the next calculation and substituting $p(\sim Fm)_n = (1-p(Fm))_n = (1-p(Fm|E)_{n-1})$, we can establish a sequence for examining an ongoing spectrum of data. Notice that we allow p(E|~Fm), (the probability of smaller objects combining in the current object given that Fm is not true) to remain at 0.5 even though the probability builds that Fm is true. This is a factor that permanently allows for Fm to eventually be found not true, which, like the Sunrise, will probably be the case at some unthinkably remote circumstance in an infinite universe.

In this manner, the Bayesian equation,

$$p(Fm|E) = \frac{p(E|Fm)*p(Fm)}{p(E|Fm)*p(Fm) + p(E|\sim Fm)*p(\sim Fm)}$$

becomes the Bayesian sequence,

$$P(Fm|E)_n = \frac{p(E|Fm)*p(Fm|E)_{n-1}}{p(E|Fm)*p(Fm|E)_{n-1} + p(E|\sim Fm)*(1-p(Fm|E)_{n-1})}$$

Starting with the initial priors plugged straight into the root from of Bayes equation we have an initial posterior probability of:

$p(Fm|E)_1 = 0.667$ (neutron constituent to the atom)
through the scalar range of $10^{-??}$ to 10^{-10} meters

From there we substitute and proceed with the sequence:

$$p(Fm|E)_1 = \frac{1 * 0.667}{(1 * 0.667) + (0.5 *(1-0.667)}$$

$p(Fm|E)_2 = 0.8$ (atom constituent to the molecule)
through the scalar range of 10^{-10} to 10^{-8} meters
$p(Fm|E)_3 = 0.888$ (molecule constituent to planet)
through the scalar range of 10^{-8} to 10^{8} meters
$p(Fm|E)_4 = 0.94$ (planet constituent to star system.)
through the scalar range of 10^{8} to10^{16} meters

$p(Fm|E)_5 = 0.969$ (star system constituent to galaxy)

through the scalar range of 10^{16} to 10^{20} meters

$p(Fm|E)_6 = 0.985$ (galaxy constituent to galaxy clusters)

through the scalar range of 10^{20} to10^{23} meters

$p(Fm|E)_7 = 0.992$ (galaxy clusters constituent to Big Bang)

through the scalar range of 0^{23} to10^{26} meters (visible universe)

$p(Fm|E)_8 = 0.996$ (Big Bang constituent to a larger structure)

through the scalar range of 10^{26} to $10^{??}$ meters

Thus, we have a consistent trend and a plausible probability that the Big Bang, in conjunction with an indeterminately broad class of Big Bang phenomena, is constituent to a larger, yet similarly finite, structure. While there are conditions in which some materials are sufficiently isolated to not be constituent to anything in particular, they are still finite and multiply manifest in the universe and might be considered in transition phase between structures.

Historical Analysis

The next range of data for our analysis of the rules of discovery is the history of our cosmological search for knowledge. The Bayesian examination of the historical sequence of the consistent divergence our cosmological models from the two rules of discovery sets up identically to the above analysis of the broad scale structure of the universe. The posterior probability, p(Ffp|E), is defined as the probability that the current cosmology fails in concise resistance of the Finite Rule and Plurality Principal, (Ffp) given the evidence (E) of the next cosmology relegating the previous cosmology to a finite, multiply manifest constituent of the new cosmology. Thus the priors are subjectively assigned as follows:

p (E|Ffp) = The probability of finding evidence E (the next cosmology proving the current cosmology to be a finite, multiply manifest constituent of the new broader cosmology) assuming that phenomenon Ffp (the current cosmology fails in concise avoidance of the two rules of discovery) is true. This value is necessarily (tautologically) 1 or true.

p (Ffp) = The prior probability of phenomenon Ffp being true *independent* of the evidence E, or *the prior probability of* Ffp. For neutrality's sake we start with a 50/50 split between Ffp (current cosmology off by rules factor) and ~Ffp (current cosmology not off by rules factor) giving Ffp = 0.5.

p (E|~Ffp) The probability of evidence E (new cosmology proving old cosmology deviated by rules factor) being true presuming that phenomenon Ffp (the current cosmology digresses by rules factor) is *not* true. We will assign p (E|~Ffp) = 0.5 permanently allowing that no matter how many times we confirm it, the fact that the new current cosmology proves the old cosmology false in confirmation of the rules of discovery there is still a 50/50 chance that could cease to be the case.

p (~Ffp) = The prior probability that Ffp is *not* true, independent of E, which is necessarily 1 – (Ffp) = p(~Ffp) = 0.5.

So we take the Flat Earth, the Ptolemaic (Earth-centered), the Copernican (Sun-centered), the "Island Universe" (Galactic), and the Big Bang cosmological models and calculate the odds that each one turned out to be a lesser, multiply manifest constituent to the subsequent model. The Flat Earth was indeed a multiply manifest constituent of a much larger Ptolemaic system, so we first calculate the root form of Bayes' Theorem for the first iteration of human cosmology:

$$p(Ffp|E)_0 = \frac{1 * 0.5}{(1 * 0.5) + (0.5 * 0.5)}$$

$p(Ffp|E)_0 = 0.667$ (Flat Earth constituent to the Ptolemaic model)

And then substitute into the sequence:

$$p(Ffp|E)_1 = \frac{1 * 0.667}{(1 * 0.667) + (0.5 *(1-0.667)}$$

*$p(Ffp|E)_1 = 0.8$ (Ptolemaic constituent to the Copernican model)

$p(Ffp|E)_2 = 0.88\underline{8}$ (Copernican constituent to the Galactic model)

$p(Ffp|E)_3 = 0.941$ (Galactic constituent to the Big Bang)

$p(Ffp|E)_4 = 0.9\underline{69}$ (Big Bang constituent to the next model)

*The Ptolemaic model does not regress from the Copernican model in as distinct a conformity to the rules of discovery as do the others, but it nonetheless represents a scalar relegation of both primacy and plurality, with the Earth being relegated in scope to a tiny constituent of the Sun's dominance and classified as one of many planets.

Convergence

An interesting property that emerges from this particular arrangement of Bayes' Theorem is that of convergence. Since in both analyses p(E|Ffp), p(E|Fm) = 1 and p(E|~Ffp), p(E|~Fm) < 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 $p(Ffp|E)_n$, $p(Fm|E)_n = 1$, as n approaches infinity no matter how pessimistically one might assign the priors.

Substituting $p(Ffp|E)_n = X_n$, $p(E|Ffp) = 1$, and $p(E|{\sim}Ffp) = \alpha$ 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 $(Ffp|E)_0$, the initial calculation of the probability in question. 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 for any phenomenon will converge to true for a large enough data set under this arrangement of Bayes' Theorem. That's not to say that the evidence will necessarily hold true, only that this property of convergence represents an additional increment of confidence in the prospect of the rules of discovery holding true in the near term.

Conclusion

Perhaps there are an insufficient number of iterations in either model to support such conclusive odds (0.996 for the structural spectrum and 0.969 for the historical spectrum, in favor) of the adherence of the Big Bang to the two rules of discovery. Perhaps a set of priors more negatively predisposed to the confirmation of the two rules of discovery would be more appropriate than the 50/50 split depicted above. But since this form of Bayes' Theorem is convergent to true, even the most negative assignment of the priors will inevitably favor of the rules of discovery, barring contrary evidence.

Due to the inherently subjective assignment of the priors, Bayes' Theorem is of more qualitative than quantitative utility. The above analysis is not performed as an accurate numerical probability assessment so much as a qualitative assessment of the clear compatibility of the rules of discovery to a legitimate Bayesian philosophical view of the universe. It is the qualitative absence of any real world deviation from the rules as much as the quantitative conformity of the two rules across the widest scale of the human experience that carries the argument. Similarly, it is the consistency with which all human cosmologies have concisely resisted these two rules and their retroactive correction to conformity under each new cosmology that best compels the provisional adoption of the rules until such point as the hard data dissuades it.

Any additions or corrections would be greatly appreciated.

Send all inquires to: author@thegodofreason.com