obits, but that the retrograde motion of the planets arises from an imposition of the motion of the earth upon them.

In the Copernican system, the appearances of planetary motion then fix many of the details. Corresponding details must be set by independent stipulation in the Ptolemaic system. The relative sizes of the planetary orbits are fixed in the Copernican system; but these sizes must be set by independent stipulation in the Ptolemaic system.[15] In the Copernican system there are only two possibilities for planets: either their mean positions align with the sun and their retrograde motions carries them to and fro across the sun; or they exhibit retrograde motion only when in opposition to the sun. This conforms with the appearances. The Ptolemic system can make no corresponding assurance. This conformity must be built in by independent supposition for each planet. These and more differences give the Copernican system a strong evidential advantage.

These last remarks are merely a sketch of a lengthy and complicated collection of inferences that demonstrate the evidential superiority of Copernican system. Laying it out in detail is challenging, especially if one is engaged in polemics. There the rhetoric calls for a compelling synopsis. How better to convey the Copernican advantage than by pointing to its simplicity and harmony in comparison with the Ptolemaic system? Yet it is simpler only in requiring fewer independent posits and more harmonious in that the determination of some features necessitates others. There is no manifestation of a deeper principle of parsimony in nature.

7. Bayes

7.1 The Problem

The examples of forms of inductive inference so far have given only qualitative assessments. If the Copernican system is better supported by the astronomical evidence than the Ptolemaic because it requires fewer independent assumptions, just how much better is that support? Merely reciting "much better" may be all we can say. To many that will fall far short of

---

[15] For an extended account, see the Chapter, "The Use of Hypotheses in Determining Distances in Our Planetary System."

what is wanted. Can we not measure support quantitatively? And if we can, might questions of strength of support be reduced to objective computations?

Something like this is the promise of Bayesian analysis. The founding tenet of objective Bayesianism is that degrees of inductive support are measured by conditional probabilities. A typical analysis begins with some prior probability distribution, which represents the support accrued by some hypothesis prior to inclusion of the evidence at issue. The import of the evidence on the inductive support of the hypothesis is found by conditionalizing on the evidence, usually through Bayes' theorem, to form the posterior probability. There is, I hope, no need to elaborate since, of all schemes in the modern literature, this one is now best known.

The difficulty with the Bayesian system is that it is too precise and irremediably so. There will be cases in which degrees of support can be represented responsibly by probabilities. They arise in narrowly prescribed problems. For example, since we can recover population frequencies for various genes, we can ask what is the probability that this sample of DNA was drawn from some donor randomly selected from the population. However evidential questions of a more foundational character are rarely given to us in a context rich in probabilities. Then insisting on a Bayesian analysis can be satisfying in the sense that we replace vague notions of strength of support by precise, numerical probabilities. However the impression of progress is an illusion. The prized numerical precision has been introduced by our own assumptions that do not reflect a corresponding precision in the system investigated. We risk mistaking our manufactured precision for that of the world.

The standard view of a Bayesian account is that probabilities are supplied by default and in abundance. The material approach reverses this. According to it, we are not authorized to any probabilities by default. Probabilities can only be introduced when the background facts warrant it; and a thorough analysis can display the pertinent warrants. Adopting that new default protects us from the spurious precision that troubles so much of Bayesian analysis. For we can only introduce precise probabilities if the precision of the facts of the context allow it. To do otherwise is to risk asserting results that are merely artefacts of applying an inductive logic ill-suited to the problem at hand.[16]

7.2 Sunrises and Laplace's Rule of Succession

The problem has been with Bayesian analysis from the outset. It can already be seen in one of the earliest Bayesian analyses. Laplace asked after the probability that the sun will rise tomorrow morning, given the past history of sunrises. This was already an established question. Before him, Hume had urged that our past history of sunrises gave no assurance of future risings. Richard Price, author of an appendix to Bayes' posthumously published paper, applied Bayes' inverse method to the problem to compute the odds of a future sunrise.[17] Laplace would now give his application of the probability calculus to the problem. His 1814 analysis (1902, p. 19) is a celebrated application of his "rule of succession." To put some formulae on Laplace's non-symbolic narrative, the analysis depended on several assumptions. We assign a probability q to the rising of the sun.

$$P(\text{rising}) = q \tag{1}$$

Antecedent to all evidence of any risings, we allow that q can have any value from 0 to 1. We represent that latitude by assign a uniform probability density p to the interval. That is,[18]

$$p(q) = 1 \qquad \text{for } 0 \leq q \leq 1 \tag{2}$$

Next Laplace assumed that the individual occurrences or otherwise of a sunrise are probabilistically independent events. These assumptions were sufficient to enable Laplace to compute the probability of a sunrise on the n+1th occasion, given a history of s risings on n past occasions:[19]

---

[16] *The Material Theory of Induction*, Ch. 10, §4 gives examples of such spurious results in the form of the inductive disjunctive fallacy ("Why is there something rather than nothing?") and the lamentable doomsday argument.

[17] For more on Hume and Price, see the chapter, "The Problem of Induction," below. See Zabell (1989) for more of the history of the rule of succession.

[18] Lest it pass unnoticed, the probability P and probability density p are distinct and should not be conflated.

[19] See the Appendix for a summary of the computation.

$$P(n+1\text{th rising} \mid s \text{ risings on } n \text{ past occasions}) = (s + 1)/(n + 2) \quad (3)$$

If the sun rose on all past n occasions, then the rule of succession gives us

$$P(n+1\text{th rising} \mid n \text{ risings on } n \text{ past occasions}) = (n + 1)/(n + 2) \quad (4)$$

The more risings we see, the better supported evidentially is the next rising. Its probability approaches one arbitrarily closely with enough risings. Laplace immediately translated this probability into a wager:

> Placing the most ancient epoch of history at five thousand years ago, or at 182623 days, and the sun having risen constantly in the interval at each revolution of twenty-four hours, it is a bet of 1826214 to one that it will rise again to-morrow.[20]

7.3 What is Wrong With It?

This precise quantitative result and its operationalization in a bet is momentarily satisfying and perhaps even thrilling, if numerical precision is the goal. Yet a moment's more reflection reveals that the precision attained is fabricated and fanciful. There are two problems, to be addressed in the next two sections:

- First, the impression of recovery of a result of some generality is illusory.
- Second, a probabilistic analysis is the wrong analysis for the problem as actually posed by Laplace.

Laplace's analysis has been chosen for scrutiny here since its simplicity enables us to see both problems quickly. We might imagine that the development of the Bayesian approach after Laplace has addressed and resolved these problems. To some extent, this has happened. Where these problems persist most notably, however, is in Bayesian analyses in philosophy of science. There these methods are routinely applied to problems with vague specifications. The goal is to supplant their vagueness with mathematical precision. This laudable goal, however, can only be achieved by imposing assumptions whose precision is unwarranted by the problems posed. As

---

[20] The computation of the number of days in 5000 years as 182623 is an obvious error, too low by a factor of 10. Five thousand years corresponds to 5,000 x 365 = 1,825,000 days or 5,000 x 365.2422 = 1,826,211 days depending on how one counts days in the year. The odds reported by Laplace of 1,826,214 to one indicate that Laplace's real estimate of the number of days in 5,000 years is 1,826,213. The erroneous 182,623 results from dropping the tens digit 1.

with Laplace's sunrises, the precision of the ensuing analysis is an illusion of our own manufacture.

7.4 Failure of Generality

Laplace's "rule of succession" is presented with a suggestion of some sort of general applicability. Perhaps it is a general demonstration that probabilistic analysis defeats Hume's skeptical challenge to inductive inference. While the application to sunrises specifically is far-fetched, perhaps it shows that probabilistic analysis can solve the sort of inductive problems Hume charged as insoluble. Or perhaps more modestly it is, at least in simple cases, a convenient starting point for how we are to think of projecting a record of successes and failures inductively into the future.

From the perspective of the material theory of induction, it does none of these. It is a theorem in probability theory, untroubling merely as a piece of mathematics. However, as an instance of inductive inference, it is untethered from real problems in the world. Any inductive rule, such as the rule of succession, can only be applied to some particular problem if the background facts of the domain warrant it. Without that tethering, it is just a piece of mathematics.

To which inductive problems can the rule be tethered? That is, which problems are such that their background facts warrant the rule. We find that there are very few and they are artificial.[21]

It is no surprise that the rule of succession fails for the real problem of sunrise prediction. The pertinent background facts are rich. Sunrises come about from the rotation of the earth on its axis; and this rotation is one that can only be disrupted by the most cataclysmic of cosmic events. Absent such a cataclysm, successive risings are perfectly correlated; and after such a cataclysm, successive failures to rise are perfectly correlated. The assumption of probabilistic independence for each sunrise fails. If we are serious about predicting such a cataclysm from, say, an errant galactic body, then our analysis would need to inquire after the distribution of such bodies in our

---

[21] We might compare this rule with the ideal gas law in the thermodynamics of gases. It is derived from highly idealized assumptions. Unlike the rule of succession, the ideal gas law applies to a wide range of ordinary gases in ordinary circumstances.

neighborhood and to hope that sufficient information is forthcoming so that probabilistic predictions of cataclysmic collisions with earth can be mounted.

Laplace had no illusions that his analysis was close to one that accommodated what we know factually of sunrises. He continued the report on the bet quoted above by saying:

> But this number is incomparably greater for him who, recognizing in the totality of phenomena the principal regulator of days and seasons, sees that nothing at the present moment can arrest the course of it.

This does not appear to be a retraction of his analysis, but may merely be a statement that it gives an excessively modest lower bound to the probability appropriate to our real epistemic situation.

If not sunrises, then might Laplace's analysis apply to the expectation of live human awakening? Then biological facts as summarized in mortality tables provide the background facts needed to assess the probability of a human awakening tomorrow, given some past history of awakenings. A 20 year old male has a 20 year history of successful awakenings. Mortality tables[22] tell us that a male has a probability of 0.998827 of surviving the next year. Taking the approximation that the probability of a successful awakening each morning in the year is the same, the probability of success on the next morning is $0.998827^{1/365} = 0.999996784$. The same computation for a 100 year old female gives us a probability of success in awakening the next morning as $0.69845^{1/365} = 0.99901722$. The rule of succession does not apply.

These examples may be multiplied. Laplace's analysis is almost never warranted by background facts. Where does it apply? Laplace's own text shows us a way. The problem of sunrises comes at the end of Laplace's Chapter 3. Virtually all the other examples in that chapter are of familiar games of chance and associated randomizers: the tossing of coins, the throwing of dice and the drawing of black or white balls randomly from an urn. Consider this problem:

> An urn contains a very large number of coins, which are biased in all possible ways. The biases are uniformly distributed over all possible values: coins with a chance of heads *q* appear in the urn with the same frequency for all *q* the entire range from 0

---

[22] Provided by the US Social Security Administration at https://www.ssa.gov/oact/STATS/table4c6.html

to 1. We select a coin at random from the urn.[23] We toss it 182623 times and find heads on every toss. What is the probability that the next toss is a heads?

It requires only a little reflection to see that all the conditions for Laplace's rule of succession are satisfied. The background facts warrant the application of Laplace's rule of succession. It assures us that the odds of a head on the next toss are 182624 to one.

Laplace's analysis illustrates a common problem with Bayesian analysis. It has a small repertoire of tractable templates. They include sampling problems, such as drawings from urns; and problems in games of chance, which are based on physical randomizers, like thrown dice, shuffled cards and tossed coins. The supposition is these templates can be applied to problems that bear only superficial resemblance to the original problems of sampling or games of chance. This supposition mostly fails. Inductive problems in the real world—especially the more interesting ones—are rarely structurally like simple problems of sampling or games of chance.

7.5 Probabilities are Inapplicable

Laplace's mention of his analysis as applying to sunrises can and, indeed, should be taken only as a colorful embellishment intended to make an arid technical problem appear less dry. For the problems is posed *by assumption* in a factually barren landscape. The problem's formulation fails to provide the background facts that are required to warrant an inductive inference. To describe the problem as inferring from the evidence of 182623 sunrises is misleading, if taken seriously. Calling them "sunrises" triggers the sorts of background knowledge mentioned above that we are supposed to discount. Successive sunrises are very strongly correlated, yet Laplace's analysis makes them probabilistically independent. A better description might be the vaguer evidence statement:

We have 182623 successes. Will the next occasion be a success?

The only answer we can give is that we cannot say. The evidence is given in a vacuity of background facts. It supports no inductive inference. We need background facts on the nature of the occurrences to warrant an inductive inference. When they are supplied, we can determine just

[23] I follow Laplace in overlooking the practical and principled difficulties of selecting randomly from an urn with an infinity (here uncountable) of balls or coins. A safer system spins a pointer on a dial to select a number randomly between 0 and 1. We then construct a coin with that number as its bias.

which inductive inferences are warranted. Which they are will vary from circumstance to circumstance. Laplace's analysis will almost never apply.

If we persist in applying a Bayesian analysis and recover results of any strength, where none are warranted, all we can conclude is that these results are artefacts of a misapplied inductive logic. Once we are alerted to the danger, it is easy to see how Bayesian analysis introduces factual presumptions under the guise of benign analytic machinery. The idea that the unspecified occurrence can be represented by a probability distribution at all is an example. It commits us to factual restrictions that go beyond the factual barrenness presumed. To assign a middling value to the probability, $P(\text{rising}) = q = 0.5$, is not to be neutral. It is to say that, loosely speaking, in situations similar to that of the analysis, we should expect an occurrence in roughly half of them.

Then there is the attempt to represent the complete openness over which value of q applies. Laplace does his best here by assuming a uniform probability distribution (2) over q. This uniform distribution once again goes beyond the factual barrenness presumed. For that distribution makes many strong claims. It says that a value of q in the interval $(0, 0.1)$ is as probable as a value of q in the interval $(0.5, 0.6)$ but only half as probable as a value of q in the interval $(0.5, 0.7)$. The interval $(0, 0.99)$ is highly probable and its complement $(0.99, 1.0)$ highly improbable. These are strong statements. The absence of background facts means that none of them are authorized.

The difficulty of representing evidential neutrality in a probabilistic analysis is well-known. Various techniques known as "imprecise probability" can be used to ameliorate the failure of a uniform probability density to represent adequately a complete indifference over the values of the parameter q.[24] In one approach, we replace the single prior probability density (2)

---

[24] Might we escape these problems by adopting subjective Bayesianism? Then the prior probability distribution is merely uninformed opinion and may be freely chosen, as long as it preserves compatibility with the probability calculus. This popular approach has had a malign effect if one's interest in inductive support and bearing of evidence. For once one allows opinion free admission into one's system, it becomes very difficult to remove its taint from one's judgments of inductive support. The limit theorems that are supposed to purge the subjectivity apply in limited, contrived circumstances that do not match the real practice of science.

over q by the set of all[25] probability densities over the interval [0, 1]. When we apply the rule of succession, instead of recovering a single probability for the next occurrence, we recover a set of probabilities. In general, there is one for each of the probability densities in the set. That we admit all probability densities gives the appearance of the requisite independence from background facts. That appearance is an illusion since we are still assuming that the probability calculus applies at all, even in weakened form. The introduction of this imprecision is fatal, however, to the recovery of a non-trivial result. For, as we see in the Appendix, the set of all prior densities includes ones that lead to all possible probabilities from zero to one for the next sunrise. We start assuming that this probability can lie anywhere between 0 and 1 and must end without any restriction on this range. We will have learned nothing from the evidence, no matter how extensive our history of sunrises may be.

7.6 Bayesian Analysis within the Material Theory of Induction

What are the prospects for Bayesian analysis from the perspective of the material theory of induction? Bayesian analyses can be applied profitably to many, specific inductive problems. Given what we know about errant galactic bodies, what should our expectations be for a cataclysmic collision with the earth that will disrupt our sunrises? Given patients with such and such prognosis, what is their life expectancy? These and many more problems like it are all welcomed by the material theory of induction. For in each case there are identifiable background facts that warrant the application of a probabilistic analysis.

Where Bayesian analysis fails is that it cannot provide an all-embracing framework with formal rules applicable to all problems of inductive inference. It will work well in specific problems, where the background facts warrant it. But any claim of general applicability, such as is sought in the philosophy of science literature, requires that the framework must be applicable to inductive problems whose background facts fail to authorize a probabilistic analysis. In these cases, persisting in applying a probabilistic analysis risks producing results that are artefacts of an inapplicable inductive logic.

---

[25] The scope of "all" is vague, but that vagueness is immaterial to the points made here. As a first pass, it designates all integrable functions with unit norm.

8. Conclusion

In reviewing the material theory of induction, this chapter has been restricted to particular instances of inductive inference. In each case, the warrant for the inferences is found in background facts. For the inference to be licit, these background facts must be truths. Since these facts make claims that commonly extend well beyond direct experience, we must ask what supports the truth of these background facts. The material theory of induction is uncompromising in its answer. The only way these facts can be supported is by further inductive inferences; and those further inductive inferences will in turn require a warrant in still further inductive inferences. How do all these inferences fit together? That is the subject of this volume and is taken up in the next chapter.

Appendix: Laplace's Rule of Succession

Consider n+1 probabilistically independent trials, each with a probability of success q, where q is itself uniformly distributed over the interval [0,1] according to (2). If there are s successes only among the first n trials, then the probability of success on the n+1th trial is given by

P = P(success on n+1th trial | s successes in first n trials)

= P(success on n+1th trial AND s successes in first n trials) / P(s successes in first n trials)

Since the number of successes s is binomially distributed, we have:

$$P = \frac{\int_0^1 q \cdot \frac{n!}{s!(n-s)!} q^s (1-q)^{n-s} p(q)\, dq}{\int_0^1 \frac{n!}{s!(n-s)!} q^s (1-q)^{n-s} p(q)\, dq} = \frac{\int_0^1 q^{s+1} (1-q)^{n-s}\, dq}{\int_0^1 q^s (1-q)^{n-s}\, dq}$$

The integrals may be evaluated using the integral identity

$$\int_0^1 q^A (1-q)^B\, dq = \frac{A!B!}{(A+B+1)!} \qquad \text{(A1)}$$

for whole numbers A and B. We recover

$$P = \frac{(s+1)!(n-s)!}{(n+2)!} \cdot \frac{(n+1)!}{s!(n-s)!} = \frac{s+1}{n+2} \qquad \text{(2)}$$

It is the rule of succession (2) of the text.

To show that alternatives to the prior probability distribution (1) can lead to $P = r$ for any r between 0 and 1, consider the family of prior probability distributions:[26]

$$p(q) = \frac{(A+B+1)!}{A!B!} q^A (1-q)^B \quad \text{where } 0 \le q \le 1$$

for A and B whole numbers. Repeating the above calculation for P, we find

$$P = \frac{\int_0^1 q^{A+s+1}(1-q)^{B+n-s}\, dq}{\int_0^1 q^{A+s}(1-q)^{B+n-s}\, dq} = \frac{(A+s+1)!(B+n-s)!}{(A+B+n+2)!} \cdot \frac{(A+B+n+1)!}{(A+s)!(B+n-s)!} = \frac{A+s+1}{A+B+n+2}$$

Rewriting P as

$$P = \frac{A}{A+B} \cdot \frac{1+(s+1)/A}{1+(n+2)/(A+B)}$$

it follows that $P \to r$ in the limit of $A, B \to \infty$ such that $A/(A+B) \to r$. That is, we can bring P arbitrarily close to any nominated $0 \le r \le 1$, merely by selecting A and B large enough in this limiting process. The prior probability $p(q)$ masses all the probability arbitrarily closely to $A/(A+B)$ in the process of taking the limit. The limit itself is no longer a function, but a distribution, the Dirac delta "function." That is

$$\lim_{\substack{A,B \to \infty \\ A/(A+B) \to r}} p(q) = \delta(q-r)$$

Selection of this distribution as a prior would force P to the value of r exactly, since all intervals of values not containing r would be assigned a zero prior probability.

References

Bacon, Francis (1620) *Novum Organum* in *Advancement of Learning: Novum Organum: New Atlantis*. *Great Books of the Western World*, Vol. 30. Chicago: University of Chicago Press, 1952.

Bartha, Paul (2010) *By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments*. Oxford: Oxford University Press.

[26] Identity (A1) assures normalization to unity.