Bayesian perspectives on mathematical practice

James Franklin

University of New South Wales, Sydney, Australia

Contents

Abstract Mathematicians often speak of conjectures as being confirmed by evidence that falls short of proof. For their own conjectures, evidence justifies further work in looking for a proof. Those conjectures of mathematics that have long resisted proof, such as the Riemann Hypothesis, have had to be considered in terms of the evidence for and against them. In recent decades, massive increases in computer power have permitted the gathering of huge amounts of numerical evidence, both for conjectures in pure mathematics and for the behavior of complex applied mathematical models and statistical algorithms. Mathematics has therefore become (among other things) an experimental science (though that has not diminished the importance of proof in the traditional style). We examine how the evaluation of evidence for conjectures works in mathematical practice. We explain the (objective) Bayesian view of probability, which gives a theoretical framework for unifying evidence evaluation in science and law as well as in mathematics. Numerical evidence in mathematics is related to the problem of induction; the occurrence of straightforward inductive reasoning in the purely logical material of pure mathematics casts light on the nature of induction as well as of mathematical reasoning.

1 Evaluating conjectures in mathematical practice

Mathematical practice cannot consist solely of someone's writing down mathematical proofs and someone else's publishing and reading them. That may be the final output (in formal pure mathematics), but there are many preliminary activities that must take place on the way. Much of involves evaluating inconclusive evidence as to whether a hoped-for theorem is true and whether one's attempts at proving it are likely to succeed.

It is similar in the art of navigation (prior to google maps), and for the same reasons. Navigation cannot consist just of paths from A to B. Back of house, there must be work to plan how to find them, to choose among alternatives, and to estimate whether a developing path is on the right track. In mathematical research too, one must navigate a path to a final result, a polished proof, and there are probabilistic skills needed to evaluate progress.

Especially in the early stages of work on a problem, mathematicians need not merely hunches and beliefs, but ones that are as a matter of fact well-supported by the evidence. The working mathematician must begin by asking questions like: Which of the many conjectures that could be generated are worth investigating? Which ones are relevant to the problem at hand? Which can be confirmed or refuted in some easy cases, so that there will be some indication of their truth in a reasonable time? Which might be capable of proof by a method in the mathematician's repertoire? Which might follow from someone else's theorem? Which are unlikely to yield an answer until after the next review of tenure? The mathematician must answer these questions to allocate time and effort. But not all answers to these questions are equally good. To stay employed as a mathematician, a proportion of these questions must be answered *well.* [Franklin 1987]]

The area where a mathematician must make the finest discriminations of this kind – and where he might, in theory, be guilty of professional negligence if he makes poor decisions – is as a supervisor advising a prospective Ph.D. student. It is usual for a student beginning a Ph.D. to choose some general field of mathematics and then to approach an expert in the field as a supervisor. The supervisor (after estimating that the student is capable of a Ph.D.) then selects a problem in that field for the student to investigate. In mathematics, more than in any other discipline, the initial choice of problem is the crucial event in the Ph.D.-gathering process. The problem must be unsolved at present, not being worked on by someone who is likely to solve it soon, but most importantly, tractable, that is, probably solvable, or at least partially solvable, by three years' work at the Ph.D. level. It is recognized that of the enormous number of unsolved problems that have been or

could be thought of, the tractable ones form a small proportion, and that it is difficult to discern which they are. The skill in rational evaluation of conjectures required of a supervisor is high.

The next stage in successful mathematical practice is applying for a grant. While it is not unknown to ask for money for already completed work, a grant application normally talks up the importance of an unsolved problem or general area, its "impact", and the applicant's prospects of making progress on it. Expert referees for grants will usually have rather fixed ideas on what problems in their fields are important, and as fellow pure mathematicians they will ignore the applicant's claims that the research will eventually lead to a cure for cancer and world peace. Thus much of the applicant's effort must go into arguing persuasively that the problem is at least largely solvable with the resources the grant will provide and that the applicant and team are the perfect people to do it. For established researchers, the correct strategy is to exhibit their track records and speak confidently of their expected timeline of results. (Several examples in Zeilberger 2012) It is more difficult in the case of early-career researchers, both for the applicant and the evaluators, as there is less track record to go on so the evaluators need to consider whether the applicant's approach is realistic in the light of both the problem and the applicant's background. They may be tempted to fall back on proxy indicators of "promise" such as "having been supervised by a Fields medallist".

None of that argument involves presenting proofs. It is probabilistic in nature but bears on the provability of theorems – and their provability with given resources, not their provability in principle.

With the grant won, or not, the next task is to work on the actual prospective theorems, developing insights and talking to fellow experts on the problem, and constantly evaluating the evidence that progress towards a solution is being made. We look at examples of evidence for several conjectures below.

It is not adequate to describe the relation of evidence to hypothesis as "subjective", "heuristic" or "pragmatic": although those descriptions may be true as far as they go. For the practice of evaluating conjectures to be successful there must be an element of what it is rational to believe on the evidence. The Bayesian perspective on evidence, to be explained below, gives a philosophical account of that.

2 The relation of conjectures to proof

Although it is not the only possibility, in the practice of pure mathematical research consideration of the probability of conjectures is usually allied to the traditional view that the main aim of pure mathematics is to construct semiformal but rigorous proofs of theorems.

As one researcher describes the process:

> The computer is often used as an exploratory tool in problem solving. This exploratory analysis is very important in gaining insight into problems and

> theorems and often leads to new conjectures. Heuristic reasoning may then lead from the conjecture to a formal theorem whose validity leads itself to further numerical investigation. When these numerical investigations present overwhelming numerical evidence that the proposed theorem is true, the researcher sets upon a search for a mathematical proof. [Schuster 1985], a more extensive account in [Wilf 2008], many examples in [Borwein, Bailey and Bailey 2004]

That is how mathematicians speak in cases such as the Riemann Hypothesis, considered below, where evidence has been collected and published over a long timespan because a famous conjecture has resisted proof.

Although purely computational evidence is not generally publishable by itself, it is possible to prove some restricted cases of a conjecture and exhibit computational evidence that it holds more generally. (e.g. [Boij 1999])]

However, there are some other possibilities for how probabilistic reasoning ca-work in mathematical practice, other than simply providing evidence towards a traditional formal or semiformal proof.

One is that there may be probabilistic proofs, that is, proofs or "proofs" that in their nature provide only a high probability that the result is true. A well-known case is probabilistic testing of primality, widely used in cryptography, where an algorithm delivers only a high probability that a number is a prime. [Rabin 1980]

In the light of such examples, there has been some discussion of whether mathematicians should abandon their model of formal deductive proof and rest content with some probabilistic substitute (e.g. [Fallis 1997], discussion in [Easwaran 2009]; related arguments in [Paseau 2015], [Sorensen 2016]). A popular article on the topic, titled "The death of proof", by the journalist John Horgan [1993] attracted some heated responses (e.g. [Krantz 2011], but some support in [Zeilberger 1993]) and by and large mathematical practice has remained committed to the deductive model of proof as the ideal. That ideal is likely to gain support from current projects on fully automated theorem provers (a progress report in [Ornes 2020]) The results may be hard for humans to survey, but they are fully deductive proofs which do not contain human errors and sloppiness.

We should also not forget the role of arguments from authority in mathematics. [Inglis & Mejía-Ramos 2009] While mathematics in principle offers the possibility of understanding any result by following its proof, it is impossible in practice to do that for all the results one must know. Mathematicians, like other people, rationally attribute high probability to the theorems signed off by respected experts. Refereeing depends on it. So does learning mathematics. (Consequences for pedagogy considered in [Aberdein 2019], section 5)

3 Applied mathematics and statistics: understanding the behavior of complex models

Even more than in pure mathematics, numerical evidence has come to the fore in the large areas of applied mathematics and computational statistics that have been opened up by the vast increases in computational power in recent decades. In such fields, one is typically dealing with a very complex mathematical object such as a simulation of climate, whose mathematical properties are well beyond the reach of analytic methods or proof. Understanding how it works – that is, its internal behavior rather than its match with the modeled outside reality – can only be approached by experimenting with its output and applying normal (probabilistic) scientific methods for drawing conclusions from experiments.

A typical area for such work is computational fluid dynamics. One begins with a complex physical phenomenon which one hopes to understand, such as turbulence generated by nonbreaking surface waves in the ocean. [Tsai, Chen and Lu 2015] One replaces the physical phenomenon with some standard numerical simulation of it, which one can run on a powerful computer with various choices of parameters. The simulation is a strictly mathematical object, whose relation to the physical phenomenon being modelled needs understanding; for one thing, typically the model is discrete and the physical phenomenon continuous (at least at the scale being considered). But experimental evidence is available that the model and phenomenon behave quite similarly, or there would be no point in the simulation. One can therefore hope to understand the modeled reality by experimenting with and analyzing the model.

The behavior of the simulation (now considered in abstraction from its role in imitating the phenomenon), is a purely mathematical question. In principle it might be addressed by solving equations and writing proofs, but almost always the simulation is too complex for that to be possible. One can only perform (numerical) experiments, that is run the simulation with different choices of parameters (such as total energy in the system and various initial conditions), and see what happens. If the phenomenon to be explained sometimes occurs and sometimes does not, progress has been made in understanding its causes. But the different simulation runs are only samples of the simulation's mathematical behavior, so there is an essentially probabilistic inference from the runs to the overall behavior, just as there is for any conclusion from finite experimental evidence to a general law.

Something similar occurs in "big data" applications where one sets up an algorithm intended to solve some statistical problem such as fraud detection or handwritten character recognition. Before testing the system "in the wild" on real data, one often seeks first to understand its performance in principle – does it have the inherent capacity to find certain structures in data? (The question is particularly urgent if there is a shortage of real data; example in [Hoffman et al 2019])That is a strictly mathematical question and so in theory subject to proof, but with a complex system of the sort typically used, proof is not attainable. So one tests it on synthetic data, that is, computer-generated data into which one has inserted struc-

tures of the kind one hopes the system will be able to detect. If the system detects them, something is known about its performance. A run with purely synthetic data is an experiment that gives rise to purely mathematical knowledge.

Or one may want to compare different available mathematical methods for a type of task, such as spatial prediction where again there is a shortage of real data. Simulated data gives a sufficient quantity of data to evaluate the capacities of the different methods. (e.g. [Harris et al 2010]) Again, the numerical experiments result in a purely mathematical conclusion.

4 The objective Bayesian perspective on evidence

There is no agreed account of what Bayesianism means, but here we present a version that is suitable as a theoretical framework for the evaluation of evidence for conjectures in pure mathematics. (Taken from [Franklin 2011] and [Franklin 2016] but based on [Keynes 1921] and [Jaynes 2003]) It is the version called (fully) objective Bayesianism or logical probability.

It holds that the relation of evidence to conclusion is a matter of strict logic, like the relation of axioms to theorems in mathematics but less conclusive. Given a fixed body of evidence – say in a trial in court, or in a dispute about a scientific theory, or in evaluating numerical evidence for a mathematical conjecture – and given a conclusion, there is a fixed degree to which the evidence supports the conclusion. If we could agree just what the standard of "proof beyond reasonable doubt" is, then, in a given trial, it is an objective matter of logical fact whether the evidence presented does or does not meet that standard, and so a jury is either right or wrong in its verdict on the evidence.

It is not essential to the Bayesian perspective that the relation of evidence to conclusion should be given a precise number. It is unlikely that in typical legal, scientific or mathematical cases a numerical probability could be calculated, even it if existed in principle. It is sufficient for objective Bayesianism that it is sometimes intuitively evident that some hypotheses, on some bodies of evidence, are highly likely, almost certain, or virtually impossible.

The most central theses of Bayesianism do not concern numbers but are certain qualitative principles of evidence. The first is the simplest principle of logical probability, called by Pólya ([1954], 4; further in [Mazur 2014]) "the fundamental inductive pattern" or "verification of a consequence". It is:

> q is a (non-trivial) consequence of hypothesis p
> q is found to be true
> So, p is more likely to be true than before

(In short, "Theories are confirmed by their consequences or predictions.") That is taken for granted as rational in scientific experiment or historical inquiry (except by strict Popperian falsificationists). The point of drawing predictions from theories and checking if they are true is normally taken to be that confirmation of predictions supports a theory. That is as applicable to mathematical conjectures as to

scientific theories: the simplest possible consequences of a general mathematical statement such as "All odd numbers have odd squares" are its instances such as "The square of 131 is odd", which can be checked by calculation. Finding the consequences to be true gives some degree of support for a generalization and can justify further work.

It is true that in some clearly defined circumstances, it is natural to associate a number with the degree to which evidence supports a conclusion. Those circumstances are cases of the "proportional syllogism" or "statistical syllogism", where the sole relevant evidence is a proportion in a set. If the sole evidence bearing on whether a certain patient will be cured of disease A by drug B is that 89% of patients with disease A are cured by drug B, then it is natural to assign a number to the probability, on that evidence, that the patient will be cured; namely, 0.89. (Even with the proportional syllogism, however, arguments with imprecise numbers are often the most applicable, such as, "The vast majority of flights arrive safely, therefore I can relax on takeoff.")

When numbers are applicable, it is usual to model the relation of evidence e to conclusion h by a number $P(h \mid e)$, between 0 and 1 inclusive, "the (logical) probability of h given e", which satisfies the usual axioms of conditional probability:

$$P(\text{not-}h \mid e) = 1 - P(h \mid e)$$

$$P(h_1 \text{ and } h_2 \mid e) = P(h_1 \mid e) \times P(h_2 \mid h_1 \text{ and } e)$$

The theorem of probability that is most useful for updating evidence is the celebrated Bayes' Theorem which expresses how the probability of evidence given a hypothesis relates to the probability of the hypothesis given the evidence:

$$P(h \mid e \ \& \ b) = P(e \mid h \ \& \ b) \times P(h \mid b) / P(e \mid b)$$

(where h is read as hypothesis, e as (new) evidence, and b as the background knowledge that is taken for granted in the context of the problem). Thus new evidence supports a hypothesis which makes the evidence likely (but only if the evidence is surprising).

While Bayes' Theorem gives its name to Bayesianism, we emphasize that for the purpose of theory evaluation in general, the numerical theorem is much less important than qualitative consequences of it such as Pólya's "fundamental inductive pattern".

Another intuitively appealing such consequence, also easily derivable, is that the more surprising a consequence of h (that is, the lower its "prior" probability $P(e \mid b)$ on the background knowledge b), the more it increases the probability of h. (Thus Einstein's surprising prediction of the gravitational bending of light gave strong support to the theory of general relativity when it was observed to be true.)

If one takes those principles to be logical, they are as applicable to evidence for conjectures in pure mathematics as they are for theories in science or allegations in courts.

The appearance in Bayes' Theorem of the "prior" probabilities $P(h \mid b)$ and $P(e \mid b)$ calls attention to the need to have some knowledge of them. For example, if one takes a "dogmatic" prior according to which a hypothesis is impossible, $P(h \mid b) = 0$, then no amount of evidence will dig one out of that hole: according the Bayes' Theorem, the "posterior" probability . $P(h \mid e \& b)$ will still be zero after any evidence e. Some logical intuition will be needed to evaluate whether a mathematical proposition is inherently unlikely or not, but such intuition seems to be available in simple cases; for example, it seems intuitively unlikely that the square of every odd number should be odd unless there was some good mathematical reason for that.

Other versions of Bayesianism are possible, such as subjective Bayesianism which requires only the axioms and does not constrain prior probabilities [Howson and Urbach 2006] and a version of objective Bayesianism that falls short of claiming that the probabilities involved in evidence evaluation are matters of strict logic. [Williamson 2010] These are less attractive in the very abstract reasoning of pure mathematics where it seems that any considerations that genuinely count can only be logical; but in any case those versions agree on the qualitative principles of probability such as Pólya's inductive pattern, which are the main concern in explaining rational methods of evaluating the strength of conjectures.

5 Evidence for and against the Riemann Hypothesis

Much of what has been said so far – both about the process of evaluating evidence for conjectures and the general principles of Bayesian reasoning – can be illustrated by the evidence that has accumulated for and against the Riemann Hypothesis, currently the most famous unsolved problem of mathematics. Because the Hypothesis has resisted proof for so long, the mathematical community has been much more explicit than usual about the evidential situation.

Expert opinion is that the Riemann Hypothesis is almost certainly true, and that a proof is not exactly imminent but may not be far off. Naturally, some of the considerations are only comprehensible to true experts, but some of them are comparatively shallow and can be widely appreciated.

Very relevant in this case are the shallowest arguments, those from authority. How people think about arguments from authority in mathematics is evident in reactions to people who claim to have proved the Riemann Hypothesis. If someone unknown presents his "result" on BBC news and at a conference run by someone who has accepted fake papers, he will not be taken seriously. [Steckles and Lawson-Perfect 2018] If the aged Fields Medallist Sir Michael Atiyah presents a sketch of a proof, it is still met with skepticism, because of the difficulty of the problem, the minimal amount of information provided, and the author's track rec-

ord of making unestablished proof claims in his old age. [Amit 2018] On the other hand when Alain Connes [2016] learnedly explains the comparative strengths of different strategies that may lead to a solution, mathematicians in general will nod in agreement, whether they understand it or not. He's the expert.

The Riemann Hypothesis states that "With certain trivial exceptions, all the (infinitely many) roots of the Riemann zeta function have real part equal to ½." For the present purpose an understanding of the Riemann zeta function as a function in the complex plane is not necessary: it is only important that the Hypothesis is a simple universal proposition like "all ravens are black." It is also true that the infinitely many non-trivial roots of the Riemann zeta function have a natural order, so that one can speak of "the first million roots". (One of many popular but informed accounts is [Derbyshire 2003])

As a straightforward generalization, the Hypothesis is susceptible to purely inductive evidence, that is, just calculating many zeros starting from the first one and checking that they have real part ½. (We consider inductive evidence in mathematics in section 7 below.) That process was begun by Riemann after he made the conjecture in 1859 and has of course been much assisted by increasing computer power. The results have been, as they developed:

Worker	Number of roots found to have real part ½
Gram (1903)	15
Backlund (1914)	79
Hutchinson (1925)	138
Titchmarch (1935/6)	1,041
Lehmer (1956)	25,000
Meller (1958)	35,337
Lehman (1966)	250,000
Rosser, Yohe & Schoenfeld (1968)	3,500,000
Brent (1979)	81,000,001
Te Riele, van de Lune et al (1986)	1,500,000,001
Gourdon (2004)	10^{13}
Platt and Trudgian [2020]	3×10^{12} with more rigor

Table 1: Number of roots of the Riemann zeta found to have real part ½

Inductions in mathematics can certainly dwarf those in science.

While this inductive evidence is generally taken to be quite strong (to some indeterminate degree), there is also a reason to deny it is close to certainty. That comes from the close connection of the Hypothesis with the prime number theorem. This theorem states that the distribution of primes thins out logarithmically, or more

exactly, the number of primes less than x is (for large x) approximately equal to the integral $\int_2^x \dfrac{dt}{\log t}$.

If tables are drawn up for the number of primes less than x and the values of this integral, for x as far as calculations can reach, then it is always found that the number of primes less than x is actually *less than* the integral. On this evidence, it was thought for many years that this was true for all x. Nevertheless Littlewood proved that this is false. While he did not produce an actual number for which it is false, it appears that the first such number is extremely large – well beyond the range of computer calculations. That gives some reason – it is very hard to say how strong – for believing that there might be a very large counterexample to the Riemann Hypothesis even though there are no small ones.

The reasons why most mathematicians' confidence in the Riemann Hypothesis does approach certainty are not taken from further numerical evidence but from its connections with other theses that are or "ought to be" right. The first of these was the prime number theorem just mentioned. It is a consequence of the Riemann Hypothesis and was proved independently in 1896. The support given by this proof is thus an instance of Pólya's schema, mentioned above, "theories are confirmed by their consequences or predictions." However, the prime number theorem is not a very strong consequence of the Riemann Hypothesis: the Riemann Hypothesis "says a lot more", so the support given by the prime number theorem is only moderate.

Two other considerations (among those accessible to ordinary mathematicians) are much stronger. One is the remarkable "Denjoy's probabilistic interpretation of the Riemann Hypothesis," an easily understood thesis in number theory which is equivalent to the Riemann Hypothesis and is independently intuitively plausible. We will not review it here (for a very brief account see [Franklin 2014], ch. 15). The other is again a result of its connection with the prime number theorem. That theorem gives a crude estimate of the density of primes (namely, the density of primes around a large number N is $1/\log N$) The Riemann Hypothesis implies a much more fine-grained knowledge of the distribution of primes (namely, that the difference between the number of primes up to x and the integral above is $O(\sqrt{x}\log x)$, the best possible.) If the Riemann Hypothesis is true, that fine-grained distribution is the simplest and most natural one, but if the Hypothesis is false, some more complicated and unnatural phenomenon must be present. The type of probabilistic reasoning involved is what Mazur [2014] calls "reasoning from randomness": that after known constraints have been taken into account, the rest should be random. (To take a simpler example, given that all primes except 2 and 5 have last digits 1, 3, 7 or 9, one can initially expect these digits to occur equally often; that is, uniformly distributed as they would be if chosen at random.)

The Riemann Hypothesis is also supported by an argument from analogy (identified by Mazur as another probabilistic argument form common in mathematics). In some famous and difficult work, André Weil proved that the analogue of Riemann's Hypothesis is true for certain other zeta functions, and his related conjec-

tures for an even more general class of zeta functions were proved to widespread applause in the 1970s. One expert says. "It seems that they provide some of the best reasons for believing that the Riemann hypothesis is true – for believing, in other words, that there is a profound and as yet uncomprehended number-theoretic phenomenon, one facet of which is that the roots ρ all lie on Re $s = \frac{1}{2}$." ([Edwards 1974], 298) However, not all experts are convinced of the closeness of the analogy and outsiders are in no position to judge.

The Riemann Hypothesis, then, is a remarkable testbed for displaying the diversity of probabilistic evidence that can bear on a mathematical conjecture.

Similar stories could be told with the other famous unsolved problems of mathematics. For example Goldbach's Conjecture is generally strongly believed to be true (on the basis of both simple numerical and other evidence) but also to be very unlikely to be proved soon. (Review in [Baker 2009], section 3.2; evidence for some other currently unsolved conjectures listed in [Aberdein 2019], section 1) It is also instructive to follow the historical development of evidence for now-proved conjectures, like Fermat's Last Theorem and the classification of finite simple groups. In those cases the growing confidence of experts both that the conjecture was true and that proof was imminent turned out to be justified. ([Franklin 2014], ch. 15) It has indeed been argued that there is a certain observable regularity in the time-to-proof of mathematical conjectures in general. [Hisano and Sornette 2013] It would be hard to deny that probabilistic evidence in mathematics is worthwhile, in the face of evidence that it has proved in the past to be a reliable guide to the (provable) truth.

6 Probabilistic relations between necessary truths?

There is one difficult question that needs consideration in applying Bayesian theory or non-deductive logic in *mathematics* in particular. $P(h \mid e)$ is intended to be a measure of the support that evidence e gives to hypothesis h. If e entails h, then $P(h \mid e)$ is 1, since h is certain given e. But in mathematics, the typical case is that e *does* entail h, though that is perhaps as yet unknown. If, however, $P(h \mid e)$ is really 1, how is it possible in the meantime to discuss the (non-deductive or probabilistic) support that e may give to h, that is, to treat $P(h \mid e)$ as less than 1? In other words, if h and e are necessarily true or false, how can $P(h \mid e)$ be other than 0 or 1?

The answer is that, in both deductive and non-deductive logic, there can be many logical relations between two propositions. Some may be known and some not. To take an artificially simple example in deductive logic, consider the argument

If all men are mortal, then this man is mortal
All men are mortal
_____

Therefore, this man is mortal

The premises entail the conclusion, certainly, but there is more to it than that. They entail the conclusion in two ways: firstly, by modus ponens, and secondly by instantiation from the second premise alone. That is, there are two logical paths from the premises to the conclusion. More realistic cases are common in mathematics, when different mathematicians produce different proofs of the same theorem, that is, different logical paths from the axioms to the theorem.

Now just as there can be two deductive paths between premises and conclusion, so there can be a deductive and non-deductive path, with only the latter known. Before the Greeks' development of deductive geometry, it was possible to argue

All equilateral (plane) triangles so far measured have been found to be equiangular
This triangle is equilateral

Therefore, this triangle is equiangular

There is a non-deductive logical relation between the premises and the conclusion: the premises inductively support the conclusion. But when deductive geometry appeared, it was found that there was also a deductive relation, since the second premise alone entails the conclusion. This discovery in no way vitiates the correctness of the previous non-deductive reasoning or casts doubt on the existence of the non-deductive relation. That relation cannot be affected by discoveries about any other relation.

So the answer to the question, "How can there be probabilistic relations between necessary truths?" is simply that those relations are additional to any deductive relations. They may be known independently of them (and are often easier to know). Once that is established, it is possible to understand Pólya's remark that non-deductive logic is better appreciated in mathematics than in the natural sciences. ([1954], vol II, 24) In mathematics there can be no confusion over natural laws, the uniformity of nature, propensities, the theory-ladenness of observation, pragmatics, scientific revolutions, the social relations of science or any other red herrings. There are only the hypothesis, the evidence and the logical relations between them.

7 The problem of induction in mathematics

As we have seen, induction, or inference from the observed to the unobserved, is found in pure mathematics as well as in the natural and social sciences. That has significance for understanding the nature of induction as well as for mathematics. If induction works in just the same way in pure mathematics as in science, its rationality and justification would seem to be independent of any contingent facts about the world, such as the obtaining of natural laws of the uniformity of nature.

Let us take an easy example, even simpler than the ones above concerning the roots of the Riemann zeta function.

The first million digits of π are random

Therefore, the second million digits of π are random

The first few digits of π are

3.14159265358979323846264338327950288419716939937510582097 4944 …

It can be seen that they are random in the sense of lacking pattern. (That notion is formalized in several ways, as in "passes standard statistical tests for randomness" or "incompressible", not "probabilistically generated", "stochastic" [Eagle 2018]; of course the digits of π are generated by a deterministic formula. For present purposes, "looking patternless" is sufficient.) It is easily checked that the first million digits of π do both look random and pass statistical tests for randomness (such as having about the same number of each digit). (Calculations now stretch to 4 trillion digits in hexadecimal [Bailey et al 2012].)

Intuitively, that is good evidence that the patternlessness will continue into the second million digits, just in the same way that observing a million black ravens is good inductive evidence that further observed ravens will be black. It is widely believed that the digits continue to be random indefinitely, although it is not proved even that there are about equal numbers of each digit. [Marsaglia 2005] (Some claims that the digits of π fail a more subtle statistical test for randomness in [Ganz 2014] with debate in [Bailey et al 2017] and [Ganz 2017].)

Gronau and Wagenmakers [2018] argue that this is a perfect case for appreciating the Bayesian perspective in pure mathematics. They apply a formal Bayesian analysis to the hypothesis that π is normal (in base 10), that is, that the ten digits appear equally often in its decimal expansion. They conclude that "After all [the first]100 million digits [of π] have been taken into account, the observed data are 1.86×10^{30} times more likely to occur under [the hypothesis that π is normal] than under [the alternative hypothesis that is it not]. The extent of this support is overwhelming." As with any Bayesian analysis with exact numbers, the final conclusion depends on a choice of prior probabilities, and the authors discuss possible choices.

It is true, as argued by Baker [2007] that there is a special problem with inductive arguments in number theory in that all the observed cases are of small numbers. Any number that can be calculated with is very small, compared to numbers in general. That bias in the evidence could raise a question as to whether any induction of the form "All observed numbers have property X, therefore all numbers have property X" could have high probability. That does not imply, however, that inductive arguments in mathematics are generally poor. Firstly, a bias in the evidence towards small numbers does not affect inductive arguments with more modest conclusions, such as "All observed numbers have property X, so the next number calculated will have property X." (For example, the argument above about the randomness of the digits of π only extrapolated a finite distance, thus keeping to small numbers.) Secondly, many other inductive arguments have a bias in the evidence, without thereby becoming worthless (though they may become less secure). For example, extrapolative inductive inference like "All observed European

swans are white, therefore all swans in the world are white" is a worthwhile inductive argument, even though the extrapolation beyond the observed range weakens it.

It is nevertheless true that purely inductive evidence is less credible in mathematics than it might be elsewhere, because of the possibility that some mathematical reason against a generalization might apply only for huge numbers. As we saw, the purely inductive evidence for the Riemann Hypothesis was a possible instance of that, and it is possible to have "overwhelming" numerical evidence for a conjecture, where the reasons for it being false also reveal why there should be such good evidence for it. [Schuster 1985] That is not the same, however, as saying that the evidence was not good in the first place or that a blanket skepticism is justified.

The occurrence of inductive arguments in pure mathematics suggests something about their nature as well as the nature of mathematics. If inductive arguments work in the necessary matter of pure mathematics, apparently on the same basis as they work in science, it seems they cannot depend on any contingent principles, such as the uniformity of nature or the holding of natural laws. The digits of π are surely the same in all possible worlds, and the rationality of arguments about them independent of any differences between possible worlds. That is more compatible with logical justifications of induction (e.g. [Stove 1986]) than ones that rely on either contingent facts like natural laws or sociological facts about what reasoners do. Again, it is easier to appreciate the force of probabilistic reasoning in mathematical cases, free of distractions.

Conclusion

The Bayesian perspective brings some order into part of the vast backstory of mathematical practice that underpins published research output. It allows us to understand the nature of the variety of methods of evaluating evidence – for pure mathematical conjectures, for the behavior of mathematical models, for the performance of statistical algorithms. Experimental evidence, checking of consequences, weighing of intuitions, arguments from analogy and expectations as to uniform distributions work the same way in mathematics as they do in natural and social sciences. The Bayesian perspective explains the unity of those methods and why they have rational force. The fact that they do work in the necessary matter of mathematics suggests that they are matters of pure logic.

References

Aberdein, A (2019) Evidence, proofs, and derivations. ZDM Mathematics Education 51, 825–834

Amit G (2018) Riemann hypothesis likely remains unsolved despite claimed proof. New Scientist 24 Sept, https://www.newscientist.com/article/2180504-riemann-hypothesis-likely-remains-unsolved-despite-claimed-proof/

Bailey DH, Borwein JM, Calude CS, Dinneen MJ, Dumitrescu M, Yee A (2012) An empirical approach to the normality of π. Experimental Mathematics 21 (4), 375–384

Bailey DH, Borwein JM, Brent RP, Reisi M (2017), Reproducibility in computational science: a case study: randomness of the digits of pi. Experimental Mathematics 26 (3) 298–305

Baker, A (2007) Is there a problem of induction for mathematics? In M. Leng, A. Paseau, & M. Potter (Eds.), Mathematical knowledge (pp. 59–73). Oxford University Press, Oxford

Baker A (2009) Non-deductive methods in mathematics, *Stanford Encyclopedia of Philosophy* http://plato.standord.edu/entries/mathematics-nondeductive

Boij, M (1999) Betti numbers of compressed level algebras. Journal of Pure and Applied Algebra 134 (2), 111–131

Borwein, M, Bailey DH, Bailey D (2004) Mathematics by Experiment: Plausible Reasoning in the 21st Century. AK Peters, Natick, MA

Connes A (2016) An essay on the Riemann Hypothesis. In JF Nash & MT Rassias (Eds.), Open Problems in Mathematics (pp. 225–257). Springer, Cham

Eagle A (2018) Chance versus randomness. Stanford Encyclopedia of Philosophy

Easwaran K (2009) Probabilistic proofs and transferability, Philosophia Mathematica 17(3), 341–362.

Edwards HM (1974) Riemann's Zeta Function. Academic, New York

Fallis, D (1997) The epistemic status of probabilistic proof. Journal of Philosophy, 94(4), 165–186

Franklin J (1987) Non-deductive logic in mathematics. British Journal for the Philosophy of Science 38(1),1–18

Franklin J (2011) The objective Bayesian conceptualisation of proof and reference class problems. Sydney Law Review 33(3) , 545–561

Franklin J (2014) An Aristotelian Realist Philosophy of Mathematics: Mathematics as the Science of Quantity and Structure. Palgrave Macmillan, Basingstoke

Franklin J (2016) Logical probability and the strength of mathematical conjectures. Mathematical Intelligencer 38 (3),14–19

Ganz RE (2014) The decimal expansion of π is not statistically random. Experimental Mathematics 23 (2), 99–104

Ganz RE (2017) Reply to "Reproducibility in computational science: a case study: randomness of the digits of Pi" [Bailey et al. 17]. Experimental Mathematics 26 (3) 306–307

Gronau QF, Wagenmakers E-J (2018) Bayesian evidence accumulation in experimental mathematics: a case study of four irrational numbers. Experimental Mathematics 27(3), 277–286

Harris P, Fotheringham AS, Crespo R, Charlton M (2010) The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. Mathematical Geosciences 42 (6), 657–680

Hisano, R, Sornette, D (2013) Challenges to the assessment of time-to-proof of mathematical conjectures. Mathematical Intelligencer 35 (4), 10–17

Hoffmann J, Bar-Sinai Y, Lee LM, Andrejevic J, Mishra S, Rubinstein SM, Rycroft CH (2019) Machine learning in a data-limited regime: Augmenting experiments with synthetic data uncovers order in crumpled sheets. Science Advances 5 (4), eaau6792

Horgan J (1993) The death of proof. Scientific American 269 (4), 92–103

Howson C and Urbach P (2006) Scientific Reasoning: The Bayesian Approach. 3rd ed. Open Court, Chicago

Inglis M, Mejía-Ramos JP (2009) The effect of authority on the persuasiveness of mathematical arguments. Cognition and Instruction, 27(1), 25–50

Jaynes ET (2003) Probability Theory: The Logic of Science. Cambridge University Press, Cambridge

Keynes JM (1921) A Treatise on Probability. Macmillan, London

Krantz SG (2011) John Horgan and "The death of proof?". In: The Proof is in the Pudding. Springer, New York, 219–222

Marsaglia G (2004) On the randomness of pi and other decimal expansions, 2005, http://www.yaroslavvb.com/papers/marsaglia-on.pdf

Mazur B (2014) Is it plausible? Mathematical Intelligencer 36 (1) 24–33

Ornes S (2020) How close are computers to automating mathematical reasoning? Quanta Magazine Aug 27, https://www.quantamagazine.org/how-close-are-computers-to-automating-mathematical-reasoning-20200827/

Paseau, A (2015) Knowledge of mathematics without proof. British Journal for the Philosophy of Science 66, 775–799

Platt D, Trudgian T (2020) The Riemann hypothesis is true up to $3 \cdot 10^{12}$, arXiv:2004.09765

Pólya G (1954) Mathematics and Plausible Reasoning (vol. I, Induction and Analogy in Mathematics, and vol. II, Patterns of Plausible Inference). Princeton University Press, Princeton

Rabin MO (1980) Probabilistic algorithm for testing primality. Journal of Number Theory 12 (1), 128–138

Schuster EF (1985) On overwhelming numerical evidence in the settling of Kinney's waiting-time conjecture. SIAM Journal on Scientific and Statistical Computing 6 (4) 977–982

Sorensen HK (2016) 2016, 'The End of Proof'? The integration of different mathematical cultures as experimental mathematics comes of age. In Larvor B (Ed). *Mathematical Cultures*, pp. 139–160. Birkhauser, Cham

Steckles K, Lawson-Perfect C (2018) Atiyah Riemann Hypothesis proof: final thoughts, aperiodical,com, Sept 28, https://aperiodical.com/2018/09/atiyah-riemann-hypothesis-proof-final-thoughts/

Stove D (1986) The Rationality of Induction. Clarendon, Oxford

Tsai W, Chen S, Lu G (2015) Numerical evidence of turbulence generated by nonbreaking surface waves. Journal of Physical Oceanography 45 (1), 174–180

Wilf HS (2008) Mathematics: and experimental science. In T Gowers, ed, Princeton Companion to Mathematics. Princeton University Press, Princeton. 991–999

Williamson J (2010) In Defence of Objective Bayesianism. Oxford University Press, Oxford

Zeilberger D (1993) Theorems for a price: tomorrow's semi-rigorous mathematical culture. Notices of the American Mathematical Society 40, No. 8, 978–981

Zeilberger D (2012) Appendix to Doron Zeilberger's Opinion 117: Links to posted Grant Proposals https://sites.math.rutgers.edu/~zeilberg/Opinion117Appendix.html