Logical probability and the strength of mathematical conjectures

James Franklin

University of New South Wales

Unlike mathematicians, statisticians fight tooth and nail about the nature of their subject. What are probabilities, really? Actual or long-run relative frequencies? Measures satisfying Kolmogorov's axioms? Physical propensities? Degrees of belief? All those views and more have their defenders, and the differences have an impact on what to believe and do on the basis of data. The frequentist Jerzy Neyman wrote an aggressive 1961 paper called 'Silver jubilee of my dispute with Fisher' [23] and debate by practitioners on issues of principle remains vigorous (e.g. [15]).

One view of probability is of special interest to mathematicians because it offers to make sense of the notion of non-conclusive evidence for conjectures in (pure) mathematics. Mathematicians certainly say things like "We have strong numerical evidence for our conjecture"; "the consensus by then was that the proof of the classification of finite simple groups was close to completion", and so on. Those forms of words do not mean anything in the deductive logic that is the official language of mathematical proof. What sense can be made of the notion of the probability that an unproved mathematical conjecture is true?

Evidence for mathematical conjectures

Since George Pólya's celebrated books of the 1950s on "mathematics and plausible reasoning" [24] there have been occasional discussions of how mathematical conjectures can be supported by non-conclusive evidence. (Many recent examples in [6]). Pólya gives an example from Euler, published in an age before it was customary for mathematicians to cover up their working evidence for conjectures once it had been

superseded by a proof. He used some daring and obviously far from rigorous methods to conclude that the infinite sum

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \ldots$$

is equal to the prima facie unlikely value $\pi^2/6$. Finding that the two expressions agreed to seven decimal places, and that a similar method of argument led to the already proved result

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \ldots = \frac{\pi}{4}$$

Euler concluded, "For our method, which may appear to some as not reliable enough, a great confirmation comes here to light. Therefore, we shall not doubt at all of the other things which are derived by the same method." Surely he was right (except perhaps for the "at all"). The evidence was not perfect, but it was very strong. He later succeeded in proving the result. ([24] vol. I, 18-21)

In modern times, there has naturally been interest in evaluating the evidence for and against the Riemann Hypothesis, while we await its proof. (Review in [12]) Expert opinion is that the Hypothesis is almost certainly true. An early piece of evidence was that it implied the Prime Number Theorem, which was proved independently. André Weil's proof of the hypothesis for analogues of the zeta function in a different context provides, says Harold Edwards in his book *Riemann's Zeta Function,* "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 = ½." ([8], 268-9) Perhaps the most substantial reason against comes again from its close connection with the Prime Number Theorem, which in one form states that the number of primes less than x is (for large x) approximately equal to the integral

$$\int\limits_{2}^{x} \frac{dt}{\log t}$$

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 that evidence, it was thought at one time that this was true for all x. Nevertheless Littlewood proved that this is false. The first number for which it is false, however, appears to be well beyond the range of computer calculations, which gives some slight hint that there may be a large counterexample to the Riemann Hypothesis even though there are no small ones.

In a recent *Mathematical Intelligencer* article [22], Barry Mazur adds to Pólya's ideas a form of non-conclusive reasoning often found in modern number theory, which he calls "reasoning from randomness". The idea is that, subject to known constraints, one should guess that everything else in the distributions of numbers is random. An older example gives the flavor. Goldbach's conjecture asserts that every even number greater than 2 is the sum of two primes. Now there are a lot of sums of two primes about, and more so as the size of numbers increases. Starting from the fact that the density of primes around the number n is about $1/\ln(n)$, an easy and rather crude statistical analysis will conclude that the number of ways of writing a large even number m as the sum of two primes should be about $m/2\ln^2 m$. Thus the "chance" of a large number failing to be the sum of two primes is very small. Of course there is no genuine chance in number theory, and it is possible (for all we know at present) for some even number to fall through the cracks and fail to be the sum of two primes. But workers in the field, as well as amateurs, rightly regard the argument as heuristically strong. Similarly, one argument in favor of the Riemann Hypothesis is that any zeros off the critical line would imply that the distribution of primes is not as random as it "ought to be".

Most of the best-known long-standing conjectures are in number theory, but efforts are made to weigh the strength of conjectures elsewhere too. The P = NP conjecture of theoretical computer science is a strange one. The majority opinion is that $P \neq NP$, but that rests largely on the rather shallow consideration that after much trying, no-one has found a polynomial-time algorithm to solve computationally-hard problems. Other than

that, comments by experts are near-random [14], while one mathematical team that offers a hint of a path to a proof via algebraic geometry estimates that their program will require a hundred years. [9]

Non-deductive reasoning occurs also in more mundane mathematical situations than that of long-standing unproved conjectures, and across all fields. 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. Anyone can generate conjectures, but which ones 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? Which problems are suitable to set as a PhD topic, as likely to be solvable or partly solvable by four years' work at the PhD level? Mathematicians must answer these questions to allocate their time and effort. But not all answers to these questions are equally good. To stay employed as mathematicians, they must answer a proportion of them *well*. But to say that some answers are better than others is to admit that some are, on the evidence at hand, more reasonable than others, that is, are rationally better supported by the evidence. That is to accept the objectivity, as a matter of logic, of non-deductive relations between evidence and hypothesis.

Despite a long tradition of interest in non-deductive reasoning, mathematicians have shown remarkably little curiosity about what could be meant by evidence for mathematical conjectures. Pólya perhaps created confusion with his word "plausible", which suggests something merely psychological, but he rightly saw that it was not a matter of subjective impressions, but with what degree of belief was *justified* by the evidence. ([24] vol. I, 68)

Who can offer an account of that objectivity? Enter logical probability, also known as non-deductive logic.

The logical probability theory of evidence

The logical theory of probability aims to explain what sort of thing evidence is. 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, or in a dispute about a scientific theory, or speculation about a mathematical conjecture – and given a conclusion, there is a fixed degree to which the evidence supports the conclusion. It says, for example, that if we could establish 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. The view was first clearly expressed in his *Treatise on Probability* of 1921 [20] by Keynes, who thought of logical probability as a degree of partial implication. A later classic work is E.T. Jaynes' *Probability Theory: The Logic of Science* [19]. (Recent explanations and defenses in [10], [11])

That view of probability contrasts with, for example:

- Psychological views that deal only in people's actual degrees of belief in propositions – defenders of logical probability are keen to emphasise the difference between what people in fact believe and what they ought to believe

- Sociological views that people's degrees of belief are socially constructed (solely) on the basis of power relations, patronage and so on

- "Subjective Bayesianism", which allows one to have any degrees of belief one likes in propositions, provided the system is "consistent", e.g. that one's degree of belief in not-p is 1 minus one's degree of belief in p. (Subjective Bayesianism led the revival of Bayesian statistics around the 1970s and 1980s, but Bayesianism has tended in a more objectivist direction since then. [5] The school of "objective Bayesianism" comes close to the logical view of probability [27])

- Frequentism and propensity interpretations of probability, which believe that all probabilities are about relative frequencies or physical propensities (biases), and that there is no such thing as what one ought to believe on non-conclusive evidence

The logical probabilist does not of course deny that there are such things as relative frequencies, stochastic processes and degrees of belief, which may satisfy appropriate axioms and arguably deserve the name "probabilities". What is claimed is that over and above all such things, there is an objective logical relation of evidence to conclusion, short of full logical implication.

It is not essential to the logical perspective that the relation of evidence to conclusion should be given a precise number, nor that it be possible to compute the logical relation between evidence and conclusion in typical cases. It is sufficient 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 logical probability do not concern numbers but are certain qualitative principles of evidence. The first is the simplest principle of logical probability, called by Pólya "the fundamental inductive pattern" ([24] vol. II, 4) 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."

Surely that is an absolutely basic principle of reasoning. How could one make progress in science or ordinary life without thinking that way? Science relies on theories being confirmed by their predictions being found true; the detective reasons that if the butler did it, the knife will be behind the sofa, and finding it there gives greater credence to his theory. The Riemann Hypothesis implies the Prime Number Theorem, so the proof of the latter supports the Hypothesis.

Imagine a tribe that did not believe in that principle of reasoning. Surmising that there are bison in the river field, they go there to hunt them. Finding none, they gain confidence in their theory and return the next day and the one after, increasing their confidence as their theory the more it is disconfirmed. It is necessary to imagine that tribe since they are not around to be interviewed. They are extinct. Confirmation of theories by their consequences is such a basic principle that, outside deductive sciences such as mathematics and logic, little reasoning is free of it.

"Theories are confirmed by their consequences" is also (as not often noted) a simple consequence of Bayes' theorem – indeed, the main content of that theorem (and thus the logical view of probability can be called a form of Bayesianism). If, by way of idealization, one did suppose that the relation of evidence to conclusion could be given a number, then it would be natural to axiomatize it by the usual axioms of conditional

probability: P($h|e$), the "logical probability of hypothesis h on evidence e", is a number between 0 and 1 inclusive satisfying

> P(not-$h|e$) = 1 - P($h|e$)
>
> P(h_1 & $h_2|e$) = P($h_1|e$) × P($h_2|h_1$ & e)

Then Bayes' Theorem, virtually a restatement of the second axiom, is:

$$P(h|e) = P(e|h) \times P(h) / P(e)$$

(where P(p) is an abbreviation, for ease of reading, of P($p|b$), b being the assumed background information.) Now if the evidence e is a consequence of the hypothesis h, then P($e|h$) = 1. So

$$P(h|e) = P(h) / P(e)$$

If e is non-trivial, that is, not known with certainty already, then P(e) < 1. So P($h|e$) equals P(h) divided by a number less than 1. So P($h|e$) > P(h), that is, the hypothesis is more probable on the evidence than it was before. In other words, theories are confirmed by their consequences.

(Exercise: Formulate, and derive from Bayes' Theorem, Pólya's "Verification of an improbable consequence" principle ([24] vol. II, 8):

> p implies q
>
> q very improbable in itself
>
> q true
>
> _____
>
> So, p very much more credible   )

Many logical probabilists and Bayesians do believe that there is at least in principle a number (between 0 and 1) expressing the (logical) probability P($h|e$) of any given hypothesis h on any given body of evidence e. Indeed, that might be called the most orthodox Bayesian position. But Keynes argued that it was impossible to give every probability a number. Certainly, it seems both impossible and pointless to debate precise numbers for, say:

P(Marilyn Monroe was murdered by the CIA | The moon is made of green cheese).
In such a case, there is no logical relation between "evidence" and conclusion, so there is no point looking for a number to express it. (Further in [11])

The proportional syllogism and prior probabilities

That is not to say that logical probabilists are shy of numbers when appropriate. Other than the "verification of a consequence" schema, the other main form of non-deductive inference is the "proportional syllogism" or "statistical syllogism", which is inherently numerical. The strength of the argument:

> Tex is a Texan
> 90% of Texans are rich
> _____
> So, Tex is rich

is very reasonably assigned a number, namely 0.9. (Note carefully that this assertion is not undermined by objections like "What if Tex were a philosopher? He'd hardly be likely to be rich then." Logical probability states the evidential relation between a hypothesis and *given* evidence, not that between the hypothesis and some other body of evidence. The fact that

$$P(\text{Tex is rich} \mid \text{Tex is a Texan and 90\% of Texans are rich}) = 0.9$$

is not called into question by $P(\text{Tex is rich} \mid e)$ being other than 0.9 for other bodies of evidence e.) Nor is there any requirement that the individual be chosen "randomly".

Reasoning from frequencies – often imprecise ones – is ubiquitous in ordinary life, in forms like "The vast majority of plane flights land safely, so I can relax on takeoff;" "most patients with disease X who take drug Y are cured, so this patient with disease X (and no contraindications) can expect to be cured with drug Y." If that pattern of reasoning were not justified, there would be no point in collecting statistics – or at least, it would be impossible to learn from them.

The success of the proportional syllogism gives some insight also into the problem of "choosing priors", which has sometimes been thought to be a fatal objection to Bayesianism. When we consider Bayes theorem in the form

$$P(h|e) = P(e|h) \times P(h) / P(e)$$

it appears that we can only get started if we already know the "prior" P(*h*), the probability of hypothesis *h* "initially", that is, on our background knowledge. If we cannot explain where that comes from, we face an infinite regress of probabilities.

The proportional syllogism shows how we can reach bedrock and achieve absolute logical probabilities. All that is required to reach the figure of 0.9 for the probability of Tex's being rich is that all Texans should count equally (as always, unless there is definite reason to think otherwise). But that is just the symmetry of individual constants, which is a principle of logic.

Mathematical applications of the proportional syllogism to conjectures seem to be rare, but not totally unknown. If we despaired of understanding anything about the inherent likelihood of proving or disproving the P = NP conjecture, and just estimated when it might be resolved by looking at the times for hard conjectures in general to be resolved – and that has been done ([1], more soundly in [18]) – that is applying a proportional syllogism. (The analyses give a 40-50% chance of the P = NP being resolved in about 15 years; betting with large stakes is not advised.)

How can there be probabilistic relations between necessary truths?

There is a puzzle concerning how there could be probabilistic relations between the necessary truths of mathematics, the resolution of which reveals something important about how logic works in mathematics.

Suppose one argued: If *e* entails *h*, then P(*h*/*e*) is 1. But in mathematics, the typical case is that *e* *does* entail *h*, though that is perhaps as yet unknown. If, however, P(*h*/*e*) is really 1, how is it possible in the meantime to discuss the (non-deductive) support that *e* may give to *h*, that is, to treat P(*h*/*e*) as not equal to 1? In other words, if *h* and *e* are necessarily true or false, how can P(*h*/*e*) be other than 0 or 1?

There must be something wrong with that reasoning, as it parallels the obviously erroneous "The coin is determined to come up either heads or tails, so the true probability of heads must be either 0 or 1." Probability is a relation between propositions, some of which are known and some not. On the evidence available before the coin throw, the probability of heads can be ½, even if on some other evidence (very precise physical measurements and calculations) it is 0 or 1.

Similarly with the evidence for mathematical propositions. There can be a deductive and non-deductive path from evidence to conclusion, with only the latter known. Before the Greeks' development of deductive geometry, it would have been 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. That 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 (and may be known independently of them).

Inductive reasoning in mathematics

Here we do not mean the (deductive) method of mathematical induction, but ordinary inductive reasoning, as in "All observed ravens have been black, therefore it is likely that the next raven to be observed will be black." Reasoning from sample to population cannot be deductive – whatever is observed in the sample, the opposite *could* happen outside the sample. In an example adapted from Bertrand Russell, the turkey's predicting it will be fed each morning, as in the past, is no guarantee against Christmas.

Philosophers have spent much effort looking for principles that might support inductive reasoning, such as the "uniformity of nature", or laws of nature, or something about the nature of causes. Does causality act as a "cement of the universe" (in the philosopher David Hume's phrase) binding the future to the past?

That line of thinking cannot be right, since simple inductive reasoning works just as well in pure mathematics. But in pure mathematics, there is no causality, no laws of nature, and mathematics is true in all universes, uniform or otherwise.

A particularly straightforward case of inductive reasoning in pure mathematics is:

The first million digits of π are random

Therefore, the second million digits of π are random.

("Random" here means the purely mathematical notion "without pattern", not "probabilistically generated", "stochastic"; the notion of patternlessness can be analyzed in several ways involving complexity and compression [7], but for the purposes of this example it can be taken to mean simply "passes standard statistical tests for randomness".)

The decimal expansion of π,

3.14159265358979323846264338327950288419716939937 …

has no apparent pattern. The first million digits have long been calculated (calculations have reached beyond one trillion). Inspection of these digits reveals no pattern, and computer calculations applying standard tests for randomness confirm this impression. It can then be argued inductively that the second million digits will likewise exhibit no pattern. This induction is a good one. Indeed, it is generally believed that the digits of π continue to be random indefinitely, though there is no proof [21]; of the simplest consequence of randomness, normality (the equidistribution of digits), recent computational scientists write "we tested positively the normality of a prefix of roughly four trillion hexadecimal digits of π … it is extraordinarily unlikely that π is not asymptotically normal base 16, given the normality of its initial segment." [2] However, an even more recent article claims that, although the calculated digits of π pass the usual suite of statistical tests, there is a more sophisticated test which the calculated digits fail "at a high level of statistical significance". [13] No doubt this unsettling result will be subject to close examination.

Many other cases of inductive reasoning in mathematics involve the calculation of large amounts of numerical evidence for long-standing conjectures. There is a distributed computing project that has checked the Goldbach Conjecture for even numbers up to and beyond 10^{18}. [16] Calculation with the Riemann Hypothesis had checked the first 10^{13} zeros of the Riemann Zeta Function by 2004. [17] They are among the biggest inductions in the world.

Actually those inductions that bear on general propositions about numbers are not very strong, taken just by themselves, for the same reason as "The roses in my garden are red, therefore all roses whatsoever are red": the extrapolation is very bold. As pointed out by Alan Baker, all numbers calculated with are small (compared to numbers in general), so it is quite a leap from what is true in a finite set of small numbers to the infinite set of (mostly) large numbers. ([3]; dramatized in [25]) The Pólya Conjecture (that the majority of numbers less than a given number have an odd number of prime factors) is often given as an example of a conjecture once believed on the basis of having no small counterexamples, but which does have large counterexamples. [26] That does not however prevent more modest inductions from being strong, such as that from the first million digits of π to the second million, or those in which purely numerical evidence is combined with other considerations, which can suggest that being small is an irrelevant attribute and thus the small numbers are a "fair" sample.

These examples illustrate Pólya's remark that non-deductive logic is better appreciated in mathematics than in the natural sciences. ([24] vol II, 24) If mathematicians could just get straight their use of non-deductive reasoning and reflect on it a little, they would have something important to contribute to the understanding of scientific reasoning.

REFERENCES

[1] S. Arbesman and R. Courtland, 2011 preview: Million-dollar mathematics problem, *New Scientist* 2972 (25 Dec 2010), 24.

[2] D.H. Bailey, J.M. Borwein, C.S. Calude, M.J. Dinneen, M. Dumitrescu and A. Yee, An empirical approach to the normality of π, *Experimental Mathematics* 21 (2014), 375-84.

[3] A. Baker, Is there a problem of induction for mathematics?, in *Mathematical Knowledge*, ed. M. Leng, A. Paseau and M.D. Potter (Oxford University Press, Oxford, 2007), 59-73.

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

[5] J. Berger, The case for objective Bayesian analysis, *Bayesian Analysis* 1 (2006), 385-402.

[6] J.M. Borwein and D. Bailey, *Mathematics by Experiment: Plausible reasoning in the 21st century* (A.K. Peters, Natick, MA, 2004).

[7] A. Eagle, Chance versus randomness, *Stanford Encyclopedia of Philosophy* (2010/2012), http://plato.stanford.edu/entries/chance-randomness/

[8] H.M. Edwards, *Riemann's Zeta Function* (Academic, New York, 1974).

[9] L. Fortnow, The status of the P versus NP problem, *Communications of the ACM* 52 (9) (2009),78-86.

[10] J. Franklin, Resurrecting logical probability, *Erkenntnis* 55 (2001), 277-305.

[11] J. Franklin, The objective Bayesian conceptualisation of proof and reference class problems, *Sydney Law Review* 33 (2011), 545-61.

[12] J. Franklin, Non-deductive logic in mathematics: the probability of conjectures, in A. Aberdein and I. Dove, eds, *The Argument of Mathematics* (Springer, Dordrecht, 2013), 11-29.

[13] R.E. Ganz, The decimal expansion of π is not statistically random, *Experimental Mathematics* 23 (2014), 99-104.

[14] W.I. Gasarch, The second P=?NP poll, *ACM SIGACT News* 43 (2) (June 2012), 53-77.

[15] A. Gelman and C.R. Shalizi, Philosophy and the practice of Bayesian statistics, *British of Mathematical and Statistical Psychology* 66 (2013), 8-38.

[16] Goldbach conjecture verification, http://sweet.ua.pt/tos/goldbach.html

[17] X. Gourdon, The 10^{13} first zeros of the Riemann Zeta Function, and zeros computation at very large height, 2004, http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeros1e13-1e24.pdf.

[18] R. Hisano and D. Sornette, Challenges to the distribution of time-to-proof of mathematical conjectures, *Mathematical Intelligencer* 35 (4) (Dec 2013), 10-17.

[19] E.T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, 2003).

[20] J.M. Keynes, *Treatise on Probability* (Macmillan, London, 1921).

[21] G. Marsaglia, On the randomness of pi and other decimal expansions, *Interstat* 5, 2005, http://www.yaroslavvb.com/papers/marsaglia-on.pdf.

[22] B. Mazur, Is it plausible*? Mathematical Intelligencer* 36 (1) (Feb 2014), 25-33.

[23] J. Neyman, Silver jubilee of my dispute with Fisher, *Journal of the Operations Research Society of Japan* 3 (1961), 145-54.

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

[25] P. Rothman and J. Rothman, Festina lente, *Mathematical Intelligencer* 20 (3) (1998), 17-18.

[26] E.W. Weisstein, Pólya Conjecture, Wolfram Mathworld, http://mathworld.wolfram.com/PolyaConjecture.html

 [27] J. Williamson*, In Defence of Objective Bayesianism* (Oxford University Press, Oxford, 2010).