

## MATHEMATICS

# Controlling our errors

Barry Mazur

**The Sato–Tate conjecture holds that the error term occurring in many major problems in number theory conforms to a specific probability distribution. That conjecture has now been proved for a large group of cases.**

Even under the best circumstances, controlling our errors is a dicey business. But as a recent series of mathematical papers<sup>1–3</sup> shows, significant strides are being made — in number theory at least. The new work amounts to a proof of a 40-year-old conjecture, known as the Sato–Tate conjecture, for a class of mathematical problems with applications in cryptography and the high-speed factorization of large numbers. This conjecture predicts the probability distribution of the error terms that pop up in these problems.

In any empirical study, errors accumulate for many reasons. All an experimentalist can hope for is to know the sources of most errors, and to be able to estimate how much trouble they cause.

The web page of the US Bureau of Transportation, for example, lists six possible causes of systematic error in the 1993 US Census count: inability to obtain information about all cases in the sample; response errors; definitional difficulties; differences in the interpretation of questions; mistakes

in coding or recoding the data obtained; and other errors of collection, response, coverage and estimation<sup>4</sup>.

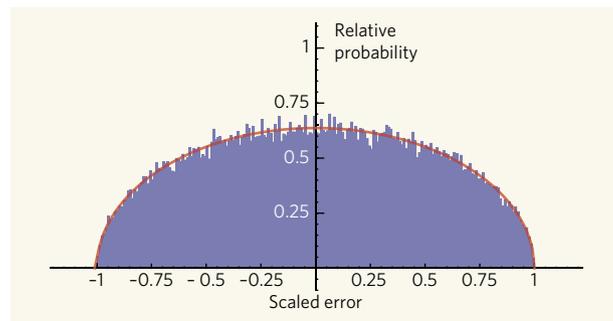
In pure mathematics, however, whenever we calculate what we hope is a good approximation to a quantity describing a phenomenon, we also have a shot at understanding — in

depth, and beyond a shadow of a doubt — the nature of the ensuing error term, defined simply as: error term = exact value – our ‘good approximation’.

Of course, if our approximation is at all good, the error term should be small. For many problems, the gold standard of goodness is what we might call square-root accuracy: that the error term scales as the square root of the quantity as it, and the phenomenon it describes, grows larger and larger. Great successes in controlling error terms in number theory were achieved in the last century. Specifically, through the work of Helmut Hasse<sup>5</sup> in the 1930s, André Weil<sup>6</sup> in the 1940s and Pierre Deligne<sup>7</sup> in the 1970s, a large class of major approximations were proved to have this kind of accuracy.

Take a (randomly chosen) example. For prime numbers  $p$ , define  $N(p)$  as the number of ways in which  $p$  can be written as a sum of 24 squares of whole numbers. (Squares of positive numbers, negative numbers and zero are all allowed.) The ordering of the squares of the numbers that occur in this summation also counts. Thus, the first prime number, 2, can already be written as a sum of 24 squares of whole numbers in 1,104 ways, because there are that many different ways in which two choices of either  $(+1)^2$  or  $(-1)^2$  can be arranged in a line where the 22 other numbers are zeros.

We know, then, that  $N(2) = 1,104$ . What about  $N(p)$  for the other prime numbers  $p = 3, 5, 7, 11, \dots$ ? A good approximation for these values  $N(p)$  turns out to be  $\frac{16}{691}(p^{11} + 1)$ . The error



**Figure 1 | Probability distribution of error terms.** The Sato–Tate distribution  $\frac{2}{\pi} \sqrt{1-x^2}$ , the smooth red curve in this figure, can be compared with the probability distribution of scaled error terms (blue bars) for the number of ways  $N(p)$  in which a prime number  $p$  can be written as a sum of 24 square numbers. The data, tabulated for primes  $p$  less than a million, agree closely with the distribution, and give hope that the Sato–Tate conjecture holds for this problem. (Courtesy of W. Stein; for details of the computation, see ref. 10.)

term on this approximation, defined as

$$\text{Error}(p) := N(p) - \frac{16}{691}(p^{11} + 1),$$

has been proven to be square-root small<sup>7</sup>. As a matter of fact,  $\text{Error}(p)$  is no larger in absolute value than  $\frac{66304}{691}\sqrt{p^{11}}$ .

This type of precision in obtaining ‘good approximations’ to difficult mathematical problems and estimating their error has led mathematicians to consider the next, but significantly deeper, tier of questions: determining the probability distribution of those error terms whose magnitudes have been shown to be square-root small.

For example, in the problem described above, let us rescale our error terms by their proven maximum order of magnitude, and ask for the probability distribution of the numbers

$$\frac{\text{Error}(p)}{\frac{66304}{691}\sqrt{p^{11}}} = \frac{N(p) - \frac{16}{691}(p^{11} + 1)}{\frac{66304}{691}\sqrt{p^{11}}}$$

as  $p$  ranges through the primes. These numbers all lie in the interval between  $-1$  and  $+1$ . In 1960, Mikio Sato (by studying numerical data) and John Tate (following a theoretical investigation)<sup>8</sup> predicted that the absolute values of the scaled error terms for data in many problems of current interest conform to a specific probability distribution; Sato and Tate shared the Wolf prize in 2003 (ref. 9). In this particular instance, this is the simple distribution curve  $\frac{2}{\pi}\sqrt{1-x^2}$ , where  $x$  ranges through the interval between  $-1$  and  $+1$  (Fig. 1). Although the Sato–Tate prediction remains unproved for this specific example, the agreement with numerical data for the first million

prime numbers gives cause for optimism.

In March this year, extraordinary strides were made towards demonstrating the truth of the Sato–Tate conjecture for one class of problems<sup>1–3</sup> related to elliptic curves. The error terms in such problems hold the key to counting solutions of algebraic equations that have a wide range of applications, including cryptography and analyses of the speed of computer algorithms.

The proof came by combining some wonderful pieces of mathematics, and the key to it is all is so-called representation theory. This branch of mathematics, in its various guises, studies abstract groups by representing them as groups of linear transformations of vector spaces. By understanding the profound number-theoretic structure behind enough of the symmetric tensor powers of a certain representation of a certain group, one can compute the probability distribution of the corresponding scaled error terms, and so confirm the Sato–Tate conjecture.

The first article of the three, by Laurent Clozel, Michael Harris and Richard Taylor (‘Automorphy for some  $l$ -adic lifts of automorphic mod  $l$  representations’)<sup>1</sup>, deals with the relationship between number theory and representation theory related to these symmetric tensor powers. Harris, Nicholas Shepherd-Baron and Taylor (in ‘Ihara’s lemma and potential automorphy’)<sup>2</sup> then show that the necessary ‘profound number-theoretic understanding’ of the  $n$ th symmetric tensor power is, for even values of  $n$ , intimately connected to the algebraic geometry of a certain beautiful family of

$(n-1)$ -dimensional vector spaces. This family, in the homogeneous variables  $X_0, X_1, X_2, \dots, X_n$ , is defined by the equation

$$X_0^{n+1} + X_1^{n+1} + X_2^{n+1} + \dots + X_n^{n+1} = tX_0X_1X_2\dots X_n,$$

where  $t$  is a parametrization variable. When  $n=2$ , this family — known as the Hessian — already played an important role in nineteenth-century geometry. The higher examples ( $n > 2$ ) are well known to theoretical physicists because of their relevance to mirror symmetry and conformal field theory, theories that are important in string theory and statistical mechanics. Finally, Taylor (in ‘Automorphy for some  $l$ -adic lifts of automorphic mod  $l$  representations II’)<sup>3</sup> overcame the last obstacle, providing a striking argument that links this ‘intimate connection’<sup>2</sup> to the truth of the Sato–Tate conjecture in the class of problems mentioned above.

This is a magnificent achievement for at least two reasons. First, the method brings synthetic unity to deep results in quite distinct mathematical fields. This coming together is as startling as the theory of continental drift that connects the shape of disparate continents.

Second, the work answers a question of delicate nature. Number theorists have long held the opinion that the ‘error terms’, despite the pejorative name, have a mesmerizingly rich structure (they are the Fourier coefficients of fascinating mathematical objects known as cusp forms) and that the keys to some of the deepest issues in their subject lie hidden in that structure. ■

Barry Mazur is in the Department of

Mathematics, Harvard University, 1 Oxford Street, Cambridge, Massachusetts 02138, USA.  
e-mail: mazur@math.harvard.edu

1. Clozel, L., Harris, M. & Taylor, R. [www.math.harvard.edu/~rtaylor](http://www.math.harvard.edu/~rtaylor) (2006).
2. Harris, M., Shepherd-Barron, N. & Taylor, R. [www.math.harvard.edu/~rtaylor](http://www.math.harvard.edu/~rtaylor) (2006).
3. Taylor, R. [www.math.harvard.edu/~rtaylor](http://www.math.harvard.edu/~rtaylor) (2006).
4. [www.bts.gov](http://www.bts.gov)
5. Hasse, H. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.* **3**, 253–262 (1933).
6. Weil, A. *Proc. Natl Acad. Sci. USA* **27**, 345–347 (1941).
7. Deligne, P. *Publ. Math. IHES* **43** (Presses Univ. France, Paris, 1974).
8. Tate, J. in *Arithmetic Algebraic Geometry* 93–110 (Harper & Row, New York, 1963).
9. Jackson, A. *Not. Am. Math. Soc.* **50**, 569–570 (2003).
10. Stein, W. <http://sage.math.washington.edu:8100/193>