



Needles and Numbers

This excursion into analytic number theory is intended to complement the approach of our textbook, which emphasizes the algebraic theory of numbers. At some points, our presentation lacks the necessary justification of convergence; but we omit these details in order to give the main spirit of the analytic approach. We begin, however, with the following curious experiment.

The Buffon Needle Experiment

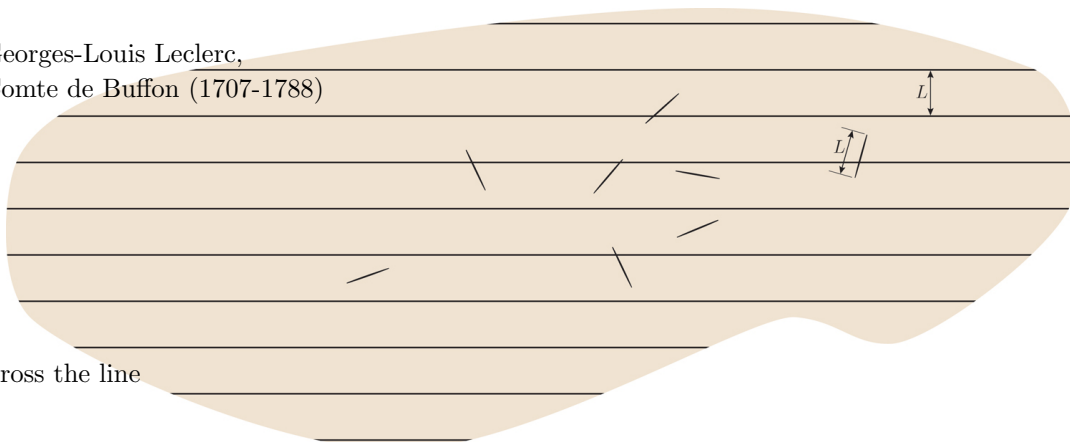
Consider a floor ruled with a family parallel lines spaced L units apart. (Imagine a hardwood floor where the floorboards have width L ; the lines are the cracks between the boards.) Now take a needle of length L and drop it on the floor. The probability that the needle crosses one of the lines is

$$Pr(\text{needle crosses line}) = \frac{2}{\pi} \approx 0.637$$

as may be shown using elementary integration (see the Appendix). Thus if the needle is dropped 1000 times (or we drop a box of 1000 toothpicks each of length L) then we expect, on average, to obtain 637 crossings. Although 637 is the *expected* number of crossings, the observed number of crossings will typically be slightly more or less than 637.



Georges-Louis Leclerc,
Comte de Buffon (1707-1788)



A typical outcome:
5 out of 8 needles cross the line

Imagine, for example, that by dropping the needle 1000 times, we obtain 641 crossings (a reasonable number to expect). If we did not know an accurate value of π , we could use this experimental data to deduce an approximate value for π , thus:

$$\frac{2}{\pi} \approx \frac{641}{1000} \quad \text{so} \quad \pi \approx \frac{2000}{641} \approx 3.12.$$

This agrees reasonably with the known value $\pi = 3.14159\dots$ to two (almost three) significant digits. If you were trying to estimate π by measuring the circumference and diameter of a coffee can, and then dividing, you probably wouldn't do any better than this! The conventional value of π is of course more reliable; but how many people can actually certify the correctness of the commonly accepted value? (Bear in mind that the digits $3.1415926535\dots$ *cannot* be verified by direct measurement of any physical circle; much more subtle formulas for π are required in order to obtain such numerical estimates for π .)

We expect that using more trials (10000, 100000, etc. instead of 1000 trials), we would obtain successively better approximations for π . This is true; yet in the limit, the accuracy of the outcome is fundamentally limited by the precision of the needle length and the spacing of the family of lines; also by the natural bias of the observer, who must decide in borderline cases whether to deem the needle as actually crossing a line.

Statistical Properties of the Distribution of Primes

Returning to number theory, imagine picking at random a number $n \in \{1, 2, 3, \dots, 10\}$. (More precisely, we mean that each of the numbers $1, 2, 3, \dots, 10$ is selected with the same probability $\frac{1}{10}$.) Clearly the probability that n is prime is

$$Pr(\text{randomly chosen } n \in \{1, 2, 3, \dots, 10\} \text{ is prime}) = \frac{4}{10} = 0.4$$

since 4 out of 10 outcomes (namely $n \in \{2, 3, 5, 7\}$) result in a prime value of n .

What if numbers are chosen randomly from a much larger interval? If we choose $n \in \{1, 2, 3, \dots, N\}$ where N is an arbitrarily large integer, then

$$Pr(\text{randomly chosen } n \in \{1, 2, 3, \dots, N\} \text{ is prime}) = \frac{\pi(N)}{N}$$

where $\pi(N)$ denotes the number of primes $\leq N$. (This is standard notation; don't confuse this ' π ' with the usual constant π used above.) This probability tends to decrease (toward 0) as N increases, since the primes become more scarce (i.e. less dense) the further we go to the right on the number line. The actual limiting density is given by the following:

Prime Number Theorem. $\lim_{N \rightarrow \infty} \frac{\pi(N)}{N/\ln N} = 1.$

This means that for large values of N , we have $\pi(N) \sim \frac{N}{\ln N}$. Put another way,

$$Pr(\text{randomly chosen } n \in \{1, 2, 3, \dots, N\} \text{ is prime}) = \frac{\pi(N)}{N} \sim \frac{N/\ln N}{N} = \frac{1}{\ln N}.$$


```

[> n:=0:
[> for k from 1 to 1000000 do
[>   if isprime(big()) then
[>     n:=n+1:
[>     fi:
[>   od:
[> n;
                                     4324

```

This says that 4324 out of 1000000 randomly selected numbers in the required range were found to be prime. So

$$Pr(\text{randomly chosen } n \in \{1, 2, 3, \dots, G\} \text{ is prime}) \approx \frac{4324}{1000000} = 0.004324,$$

as compared with the estimate

$$\frac{1}{\ln G} = \frac{1}{\ln(10^{100})} = \frac{1}{100 \ln 10} \approx 0.00434.$$

Not bad.

Next: Randomly select $m, n \in \{1, 2, 3, \dots, N\}$. (We intend that each possible value $n \in \{1, 2, 3, \dots, N\}$ is selected with the same probability $\frac{1}{N}$, and the same for m ; moreover, choices of m are independent of choices of n . In statistical language, m and n are independent identically distributed variables with a discrete uniform distribution.) What is the probability that m and n are relatively prime? Interestingly, if N grows large, this probability approaches neither 0 nor 1, but rather approaches a value $\frac{6}{\pi^2} \approx 0.6079 \approx 61\%$:

$$\lim_{N \rightarrow \infty} Pr(\text{randomly chosen } m, n \in \{1, 2, 3, \dots, N\} \text{ are relatively prime}) = \frac{6}{\pi^2} \approx 0.6079.$$

Let's check this using MAPLE to randomly select a million pairs of numbers $m, n \in \{1, 2, 3, \dots, G\}$ and count how many relatively prime pairs are obtained:

```

[> n:=0:
[> for i from 1 to 1000000 do
[>   if gcd(big(),big())=1 then
[>     n:=n+1:
[>     fi:
[>   od:
[> n;
                                     607710

```

Since 607710 out of 1000000 pairs are relatively prime,

$$Pr(\text{randomly chosen } m, n \in \{1, 2, 3, \dots, G\} \text{ are relatively prime}) \approx \frac{607710}{1000000} = 0.60771$$

which compares well with the estimate above. Or if we don't take the value of π as given, we could *estimate* the value of π using our experimental data:

$$\frac{6}{\pi^2} \approx 0.60771 \quad \text{so} \quad \pi \approx \sqrt{\frac{6}{0.60771}} \approx 3.1421.$$

Actually, our calculations were based on the limiting probability as $N \rightarrow \infty$; since G is finite, our probability is not quite $\frac{6}{\pi^2}$, which throws off our estimate of π . But you get the idea. This is all reminiscent of the Buffon Needle Experiment; but why is the probability that two large numbers are relatively prime, related to π ? I dare you to show me a circle in this computation. Actually the explanation lies in the value $\zeta(2) = \frac{\pi^2}{6}$ of the Riemann zeta function, as we now describe.

The Riemann Zeta Function

First let's recall that the Riemann zeta function is defined by

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} = 1 + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \frac{1}{5^x} + \frac{1}{6^x} + \dots$$

for all $x > 1$. This function was the subject of a famous 1859 paper by Riemann, but was

Bernhard Riemann
(1826-1866)



Leonhard Euler
(1707-1783)

earlier investigated by Leonhard Euler who obtained the following factorization:

$$\zeta(x) = \prod_{p \text{ prime}} \left(\frac{1}{1 - \frac{1}{p^x}} \right) = \left(\frac{1}{1 - \frac{1}{2^x}} \right) \left(\frac{1}{1 - \frac{1}{3^x}} \right) \left(\frac{1}{1 - \frac{1}{5^x}} \right) \left(\frac{1}{1 - \frac{1}{7^x}} \right) \left(\frac{1}{1 - \frac{1}{11^x}} \right) \times \dots$$

To verify that the *Euler factorization* of $\zeta(x)$ holds, expand the right hand side using $\frac{1}{1-t} = 1 + t + t^2 + t^3 + t^4 + \dots$, valid for $-1 < t < 1$, to obtain

$$\begin{aligned} \prod_{p \text{ prime}} \left(\frac{1}{1 - \frac{1}{p^x}} \right) &= \left(1 + \frac{1}{2^x} + \frac{1}{4^x} + \frac{1}{8^x} + \dots \right) \left(1 + \frac{1}{3^x} + \frac{1}{9^x} + \dots \right) \left(1 + \frac{1}{5^x} + \dots \right) \\ &\quad \times \left(1 + \frac{1}{7^x} + \dots \right) \left(1 + \frac{1}{11^x} + \dots \right) \left(1 + \frac{1}{13^x} + \dots \right) \times \dots \\ &= 1 + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \frac{1}{5^x} + \frac{1}{6^x} + \frac{1}{7^x} + \frac{1}{8^x} + \frac{1}{9^x} + \frac{1}{10^x} \\ &\quad + \frac{1}{11^x} + \frac{1}{12^x} + \frac{1}{13^x} + \frac{1}{14^x} + \frac{1}{15^x} + \dots \end{aligned}$$

Here I have expanded, then listed terms in decreasing order of size, starting with the largest term 1. The fact that each term $\frac{1}{n^x}$ (for $n \geq 1$) appears in this expansion, is due to the fact that every positive integer n is expressible as a product of prime powers. The fact that each term $\frac{1}{n^x}$ occurs *just once*, is due to the fact that the prime factorization of n is unique. We see that the Euler factorization of $\zeta(x)$ is a consequence of the Fundamental Theorem of Arithmetic (the theorem asserting the existence and uniqueness of prime factorizations for the positive integers). Conversely, we can deduce the Fundamental Theorem of Arithmetic from the Euler factorization of $\zeta(x)$. Simply put, the Euler factorization of $\zeta(x)$ is a restatement of the Fundamental Theorem of Arithmetic. The factor $\left(\frac{1}{1-\frac{1}{p^x}}\right)$ is called the *Euler factor* for the prime p .

The fact that the series $1 + \frac{1}{2^x} + \frac{1}{3^x} + \frac{1}{4^x} + \dots$ converges for $x > 1$ and diverges for $x \leq 1$, is well-known to Calculus II students; this follows from the Integral Test since $\int_1^\infty \frac{1}{t^x} dt$ converges for $x > 1$ and diverges for $x \leq 1$. However, the Riemann zeta function $\zeta(x)$ itself is defined for all $x \neq 1$. (Compare with the expansion $\frac{1}{1-t} = 1 + t + t^2 + t^3 + \dots$ where the series converges only for $-1 < t < 1$, whereas the function $\frac{1}{1-t}$ is defined for all $x \neq 1$.) Indeed, the Riemann zeta function $\zeta(s)$ is defined for *all complex* values $s \neq 1$. Writing $s = x + iy$ where $x, y \in \mathbb{R}$, the series and product expansions for $\zeta(s)$ converge only in the right half-plane $x > 1$, whereas the function $\zeta(x + iy)$ itself is defined whenever $(x, y) \neq (1, 0)$. The *big* question is, for which complex values of s does $\zeta(s) = 0$? Aside from trivial solutions where s is a negative real number, it is conjectured that all these solutions lie on the vertical line $x = \frac{1}{2}$ in the complex plane. This was suggested by Riemann in his famous 1859 paper, and so it has become known as the *Riemann Hypothesis*. This question (RH) is generally recognized as the most significant unsolved problem in mathematics. Before describing the significance of RH, let's observe some more accessible connections between $\zeta(s)$ and the distribution of primes.

Suppose that N is exceedingly large. In order for two randomly chosen numbers $m, n \in \{1, 2, 3, \dots, N\}$ to be relatively prime, it is necessary and sufficient that they have no prime factor in common. Now for each prime p , the number of elements in $\{1, 2, 3, \dots, N\}$ divisible by p is approximately $\frac{N}{p}$. (This value is *exact* if N is divisible by p ; otherwise, it is a very good *approximation*.) The probability that m is divisible by p is approximately $\frac{N/p}{N} = \frac{1}{p}$. Also, the probability that n is divisible by p is approximately $\frac{1}{p}$. The probability that *both* m and n are divisible by p is therefore approximately

$$\frac{1}{p} \times \frac{1}{p} = \frac{1}{p^2}.$$

(The reason for multiplying the two probabilities, is that m and n are chosen independently.) Thus the probability that p is *not* a common factor of m and n is

$$1 - \frac{1}{p^2}.$$

This is true for *every* prime p . Now if p and q are *distinct* primes, then divisibility by p is independent of divisibility by q . Thus, the probability that m and n do not have p as a common factor, *and* m and n do not have q as a common factor, is

$$\left(1 - \frac{1}{p^2}\right)\left(1 - \frac{1}{q^2}\right).$$

(Once again, the justification for multiplying is the independence of divisibility by p and by q .) This same argument can be applied to all possible prime divisors; thus the probability that *no* prime p is a common divisor of m and n is

$$\prod_{p \text{ prime}} \left(1 - \frac{1}{p^2}\right) = \left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{5^2}\right)\left(1 - \frac{1}{7^2}\right)\left(1 - \frac{1}{11^2}\right) \times \cdots = \frac{1}{\zeta(2)}$$

using the Euler factorization of $\zeta(x)$. All that remains is to determine the value of $\zeta(2)$.

Euler's Determination of $\zeta(2)$

Many proofs of the identity $\zeta(2) = \frac{\pi^2}{6}$ are now available. Some of these proofs are more elementary than others. Some proofs are elementary but rather technical. We choose to present Euler's determination of $\zeta(2)$, which is both elementary and nontechnical. This argument does lack rigor, but all the missing details can be supplied using a course in complex analysis.

Consider the function $f(z) = \sin z$. Since $f(\pi) = 0$, the function $f(z)$ should have $\left(1 - \frac{z}{\pi}\right)$ as a factor. This factor *cannot appear twice* in $f(z)$; for if $f(z) = \left(1 - \frac{z}{\pi}\right)^2 g(z)$ then

$$f'(z) = \cos z = \left(1 - \frac{z}{\pi}\right)^2 g'(z) + 2\left(1 - \frac{z}{\pi}\right)g(z)$$

and then evaluating at $z = \pi$ gives $-1 = 0$, which is impossible. So in fact $f(z) = \left(1 - \frac{z}{\pi}\right)g(z)$ where $g(z)$ does not have any factor of the form $\left(1 - \frac{z}{\pi}\right)$. Similar reasoning shows that each of the factors

$$z, \left(1 - \frac{z}{\pi}\right), \left(1 + \frac{z}{\pi}\right), \left(1 - \frac{z}{2\pi}\right), \left(1 + \frac{z}{2\pi}\right), \left(1 - \frac{z}{3\pi}\right), \left(1 + \frac{z}{3\pi}\right), \dots$$

should appear in $f(z)$ exactly once. And there are no more factors in $f(z)$, since we have taken care of all the places where $f(z)$ has roots. This means that

$$f(z) = Cz \left(1 - \frac{z}{\pi}\right) \left(1 + \frac{z}{\pi}\right) \left(1 - \frac{z}{2\pi}\right) \left(1 + \frac{z}{2\pi}\right) \left(1 - \frac{z}{3\pi}\right) \left(1 + \frac{z}{3\pi}\right) \times \cdots$$

where C is a constant. (Why not allow a function $C(z)$ instead of simply a constant C ? Such a function $C(z)$ would have to be reasonably smooth, and not have any roots; and it would have to satisfy certain growth constraints. Using complex analysis, one can deduce from this that $C(z) = C$ is a constant. We omit the details of this argument.) Now expanding gives

$$\begin{aligned} f(z) &= Cz \left(1 - \frac{z^2}{\pi^2}\right) \left(1 - \frac{z^2}{4\pi^2}\right) \left(1 - \frac{z^2}{9\pi^2}\right) \left(1 - \frac{z^2}{16\pi^2}\right) \left(1 - \frac{z^2}{25\pi^2}\right) \times \cdots \\ &= Cz - \frac{C}{\pi^2} \left(1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \cdots\right) z^3 + (*)z^5 + (*)z^7 + \cdots \end{aligned}$$

where the coefficients $(*)$ are not needed at this moment. However, Euler knew that

$$\sin z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots = z - \frac{z^3}{6} + \frac{z^5}{120} - \frac{z^7}{5040} + \cdots$$

for all $z \in \mathbb{C}$. (Our Calculus II students also know this in the case z is real.) Equating coefficients of z and z^3 from our two series for $f(z) = \sin z$ gives

$$C = 1; \quad \frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \cdots = \zeta(2).$$

Although Euler did not have the tools necessary to justify all the steps in this argument, he realized that the predicted value $\zeta(2) = \frac{\pi^2}{6} \approx 1.64493\dots$ was probably correct because it was in agreement with numerical estimates for the sum $\sum_{n=1}^{\infty} \frac{1}{n^2}$ using partial sums.

Other values of $\zeta(k)$

We extend this argument from a pair of numbers m, n to a set of numbers m_1, m_2, \dots, m_k . If m_1, m_2, \dots, m_k are integers, not all zero, we define $\gcd(m_1, m_2, \dots, m_k)$ to be the largest integer dividing all of the numbers m_1, m_2, \dots, m_k . We say that m_1, m_2, \dots, m_k are *relatively prime* if their greatest common divisor is 1. For example, 10, 12, 15 are relatively prime, even though *no two* of these numbers are relatively prime. The same argument as above shows that for k numbers m_1, m_2, \dots, m_k chosen randomly from $\{1, 2, 3, \dots, N\}$, the probability that $\gcd(m_1, m_2, \dots, m_k) = 1$ tends to $\frac{1}{\zeta(k)}$ as $N \rightarrow \infty$. Here I am assuming $k \geq 2$; but for $k = 1$, there is a natural interpretation of the result that also holds, as follows. For a single positive integer m , the greatest divisor of m is m itself, i.e. $\gcd(m) = m$;

and for all but one positive value of m , we have $\gcd(m) > 1$. So the probability that a single number m is ‘relatively prime’ is zero. If we interpret $\zeta(1) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots = \infty$ (the divergence of the harmonic series), then $\frac{1}{\zeta(1)} = 0$ agrees with the probability of zero for a single number to be relatively prime.

For even integer values of k , nice formulas are available for $\zeta(k)$; for example,

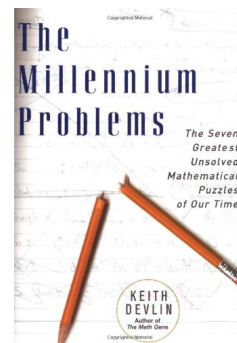
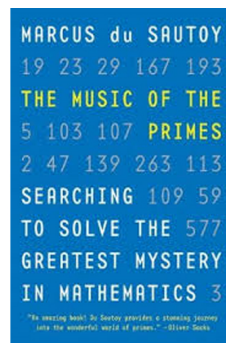
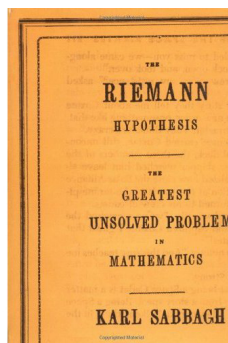
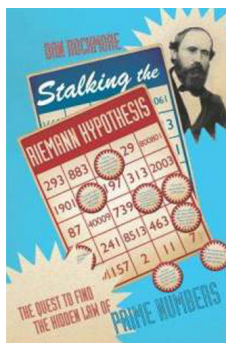
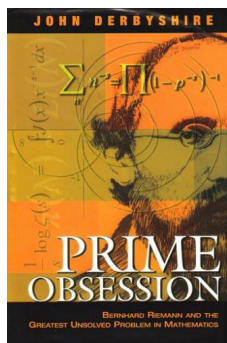
$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945}, \quad \zeta(8) = \frac{\pi^8}{9450}, \dots$$

There is a general formula for $\zeta(2r)$, which however involves Bernoulli numbers. No explicit formula is known for $\zeta(3), \zeta(5), \zeta(7)$, etc. In 1979, Apéry proved that $\zeta(3)$ is irrational; but beyond this, almost nothing is known.

The importance of the Riemann Hypothesis and zeta function

Several popular books have appeared in recent years dealing with $\zeta(s)$ and RH, including

- J. Derbyshire, *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*, Joseph Henry Press, Washington D.C., 2003.
- D. Rockmore, *Stalking the Riemann Hypothesis: The quest to find the hidden law of prime numbers*, Vintage Books, New York, 2005.
- K. Sabbagh, *The Riemann Hypothesis: The Greatest Unsolved Problem in Mathematics*, 2nd ed., Farrar, Straus and Giroux, 2004.
- M. du Sautoy, *The Music of the Primes: Searching to Solve the Greatest Mystery in Mathematics*, HarperCollins, 2003.
- K. Devlin, *The Millennium Problems: The Seven Greatest Unsolved Mathematical Puzzles of Our Time*, Basic Books, 2002.



Our discussion of the probability that a set of numbers is relatively prime, provides only a mere hint of the importance of the Riemann zeta function $\zeta(s)$. Much more generally,

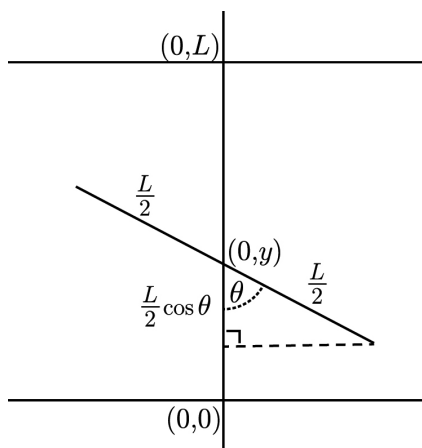
questions regarding any statistical properties of the distribution of primes involve aspects of $\zeta(s)$. In particular, since the Prime Number Theorem gives the approximation

$$\pi(N) = \frac{N}{\ln N} + \varepsilon_N$$

where the ‘error term’ ε_N is small compared to the dominant term $\frac{N}{\ln N}$, namely $\frac{\varepsilon_N}{N/\ln N} \rightarrow 0$ as $N \rightarrow \infty$, we see that any better estimates for $\pi(N)$ require a better understanding of the size of the error term ε_N . It is known that in fact ε_N has size roughly $\leq \frac{N}{(\ln N)^2}$; but the Riemann Hypothesis would yield better bounds for ε_N . This observation gives a better sense of the significance of the Riemann Hypothesis.

Appendix: Explanation of the Buffon Needle Experiment

Finally, let’s explain the probability of crossing a line in Buffon’s needle experiment. We may assume that the ruling lines on the floor are the lines $y = 0$, $y = \pm L$, $y = \pm 2L$, $y = \pm 3L$, etc. Consider a randomly dropped needle, as shown. Using symmetry, there is no loss of generality in assuming that the center of the needle lies at a point $(0, y)$ with $0 \leq y \leq \frac{L}{2}$, and that the needle forms an angle $\theta \in [0, \frac{\pi}{2}]$ with respect to the y -axis as



shown. For each center $(0, y)$, the needle crosses the x -axis iff $\cos \theta \in [\frac{2y}{L}, 1]$, iff $\theta \in [0, \cos^{-1} \frac{2y}{L}]$, which occurs with probability $\frac{\cos^{-1}(2y/L)}{\pi/2}$. Averaging over all $y \in [0, \frac{L}{2}]$ gives an overall probability of

$$\frac{1}{L/2} \int_0^{L/2} \frac{2}{\pi} \cos^{-1} \left(\frac{2y}{L} \right) dy = \frac{2}{\pi} \int_0^1 \cos^{-1} u du = \frac{2}{\pi}$$

for the needle to cross the x -axis.