Harald Cramér and the distribution of prime numbers*

AND

"It is evident that the primes are randomly distributed but, unfortunately, we don't know what 'random' means." — R. C. Vaughan (February 1990).

After the first world war, Cramér began studying the distribution of prime numbers, guided by Riesz and Mittag-Leffler. His works then, and later in the mid-thirties, have had a profound influence on the way mathematicians think about the distribution of prime numbers. In this article, we shall focus on how Cramér's ideas have directed and motivated research ever since.

One can only fully appreciate the significance of Cramér's contributions by viewing his work in the appropriate historical context. We shall begin our discussion with the ideas of the ancient Greeks, Euclid and Eratosthenes. Then we leap in time to the nineteenth century, to the computations and heuristics of Legendre and Gauss, the extraordinarily analytic insights of Dirichlet and Riemann, and the crowning glory of these ideas, the proof the "Prime Number Theorem" by Hadamard and de la Vallée Poussin in 1896.

We pick up again in the 1920's with the questions asked by Hardy and Littlewood, and indeed by Cramér. We shall see how their legacy has influenced research for most of the rest of the century, particularly through the 'schools' of Selberg, and of Erdös, and with the "large sieve" in the sixties. Then the eighties when the hitherto seemingly solid bedrock of heuristic and conjecture was shattered by a short, brilliant paper of Maier; and now, the nineties, when we are picking up the pieces, trying to make sense of what we now know.

Let's start with the one mathematical proof that every mathematician and statistician knows, *Euclid's proof of the infinitude of primes*:

Suppose, on the contrary, that there are only finitely many, call them $p_1, p_2, \dots, p_r$. *Suppose that* q *is a prime factor of the integer* $p_1 \dots p_r + 1$. *Evidently* q *must be in amongst the list* $p_1, p_2, \dots, p_r$, *say* $q = p_j$, *so that* q *divides both* $p_1 \dots p_r$ *and* $p_1 \dots p_r + 1$. *Therefore* q *must divide their difference, 1, which is impossible.*

The other great contribution to our thinking about prime numbers, dating back to Greek times, came from Eratosthenes. He showed how to create a list of all of the primes up to x, simply by knowing all of the primes up to $\sqrt{x}$. His idea was to write down all numbers up to x, then cross out every 2nd number, then every 3rd number, then every 5th number, ... and indeed every p^{th} number for each prime $p \le \sqrt{x}$. Once one has finished doing that, the numbers that are not crossed out (or '*sieved*') are the primes between $\sqrt{x}$ and x; and one can repeat this algorithm to then get the primes between x and x^2, and then between x^2 and x^4, et cetera.

As an example we find the primes between 5 and 30:

~~2~~	~~3~~	~~4~~	~~5~~	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~

The Sieve of Eratosthenes: Sieving with 2, 3 and 5.

We can even try to guess at how many primes there are up to x, using the 'Sieve of Eratosthenes' as a guide:

Since there are $x/2$ even integers up to x within an error of ± 1, the number of integers left after we sieve out those divisible by 2 is $\approx x - x/2 \approx x/2$. Similarly, of the $x/2$ remaining integers, approximately one-third are divisible by 3, and so the number remaining after sieving by both 2 and 3 is $\approx \left(1-\frac{1}{3}\right) x/2 = \left(1-\frac{1}{3}\right)\left(1-\frac{1}{2}\right) x$; or, more precisely, $\left(1-\frac{1}{3}\right)\left(1-\frac{1}{2}\right) x$ within an error of ± 2. Continuing in this way, if we sieve out the integers up to x by the k primes $\le y$, then there will be $\approx \prod_{p\le y}\left(1-\frac{1}{p}\right)\cdot x$ integers remaining; in fact within an error of $\pm 2^{k-1}$. Therefore, if we sieve out with all of the primes $p \le \sqrt{x}$ (which we know will leave precisely the prime numbers between $\sqrt{x}$ and x) then we expect to have

$$\prod_{p\le\sqrt{x}}\left(1-\frac{1}{p}\right)\cdot x$$

integers left. However, this is by no means guaranteed since the potential error term $\pm 2^{k-1}$ is by now far, far larger than the main term. It seems plausible that these 'error terms' do not accumulate in a surprising way, in that sieving by different primes may be thought of as essentially independent events, and so we might expect the number of integers left to be quite close to the estimate guessed at above.

In 1874, Mertens showed that the product above obeys the asymptotic formula

$$\prod_{p\le y}\left(1-\frac{1}{p}\right) \sim \frac{\mathrm{e}^{-\gamma}}{\log y} \tag{1}$$

where γ, the Euler-Mascheroni constant, is given by

$$\gamma = \lim_{n\to\infty}\left\{1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\log n\right\}.$$

Therefore, our heuristic here, based on the notion that sieving by the different primes $\le \sqrt{x}$ are independent events, implies that we expect there to be

$$\sim 2\mathrm{e}^{-\gamma} x/\log x \tag{2}$$

primes $\le x$, where $2\mathrm{e}^{-\gamma} \approx 1.12292\ldots$[1].

At this point we digress to give Euler's proof of the infinitude of primes (1793), partly because of its influence on the later sensational work of Dirichlet and Riemann, and partly because of its elegance and simplicity. Euler's idea was to use an identity based on what we now call the Fundamental Theorem of Arithmetic, that is that every positive integer has a unique way of being written as a product of

[1] A constant which does not seem to have a simpler definition, and seems likely to be transcendental.

prime numbers. Since any product of prime numbers is evidently a positive integer, we get the following identity: For any s with $\mathrm{Re}(s) > 1$, we have

$$\begin{aligned}\sum_{n\geq 1}\frac{1}{n^s} &= \sum_{a_2,a_3,a_5,\cdots\geq 0}\frac{1}{2^{a_2 s}3^{a_3 s}5^{a_5 s}\cdots}\\ &= \left(\sum_{a_2\geq 0}\frac{1}{2^{a_2 s}}\right)\cdot\left(\sum_{a_2\geq 0}\frac{1}{3^{a_3 s}}\right)\cdot\left(\sum_{a_5\geq 0}\frac{1}{5^{a_5 s}}\right)\cdots\\ &= \prod_{p \text{ prime}}\left(\sum_{a\geq 0}\frac{1}{p^{as}}\right) = \prod_{p \text{ prime}}\left(1-\frac{1}{p^s}\right)^{-1}. \qquad (3)\end{aligned}$$

Taking the limit as $s \to 1$ along the real line, from above, we see that

$$\prod_{p \text{ prime}}\left(1-\frac{1}{p}\right) = \sum_{n\geq 1}\frac{1}{n} = 0. \qquad (4)$$

This not only establishes that there are infinitely many primes, as promised, but also gives some hint as to how many primes there are up to x: Since (4) implies that

$$\sum_{p \text{ prime}}\frac{1}{p} \quad \text{diverges},$$

we might suppose that the primes are 'more numerous' than any sequence of integers for which the corresponding sum converges. For example, $\sum_{n\geq 1}\frac{1}{n^2}$ converges and so the primes are, in an appropriate sense, 'more numerous' than the squares. A similar argument, based on the fact that $\sum_{n\geq 1}\frac{1}{n^s}$ converges if $s > 1$ implies that $\pi(x)$, the number of primes up to x, is $> x^{1-\varepsilon}$ if x is sufficiently large, for any fixed $\varepsilon > 0$.

If we return now to the sieve of Eratosthenes, but instead sieve with only the k primes up to some fixed y, then we immediately get the upper bound

$$\pi(x) - \pi(y) \leq \prod_{p\leq y}\left(1-\frac{1}{p}\right)\cdot x + 2^y,$$

where the "2^y" bounds the accumulation of error terms. Letting $y \to \infty$, though at a rate far slower than x (for instance taking $y = \log x$), we find that $\pi(x) = o(x)$; in fact $\pi(x) = O(x/\log\log x)$.

Mathematicians (for instance, Legendre) had for some time recognized the importance of proving the analogue of Euclid's Theorem for arithmetic progressions: That if the greatest common divisor of a and q is 1 (written $(a,q) = 1$) then there are infinitely many primes $\equiv a \pmod q$[2]. In 1837 Dirichlet modified Euler's identity (3) appropriately, and managed to solve this difficult question. The proof is beautifully explained in Davenport's book. We shall just comment that an essential (and extremely surprising) ingredient in Dirichlet's proof is a link made between the value at $s = 1$ of a complicated analytic function[3] and the structure of a group that appears when describing the algebra of binary quadratic forms. In that we see the principle that guides much of modern number theory and it can be said that the Main Conjecture of Iwasawa Theory, the Birch-Swinnerton Dyer Conjectures,

[2] that is in the arithmetic progression $a, a+q, a+2q, \ldots$

[3] a 'relation' of the function in (3)

as well as the Taniyama-Shimura-Weil Conjecture[4], all propose suitable analogues of Dirichlet's formula.

Gauss, at the very end of the eighteenth century and Legendre, in the early part of the nineteenth century, considered the question of estimating $\pi(x)$, the number of primes up to x. Gauss never published his work, but as an old man, wrote in a letter to Encke, on Christmas eve 1849, (liberally translated)

> *"As a boy I considered the problem of how many primes there are up to a given point. From my computations, I determined that the density of primes around x, is about $1/\log x$."*[5]

One can interpret this as a statement of probability in order to guess at a value for $\pi(x)$: Assume that an integer n, 'close' to x, is prime with 'probability' $1/\log x$. Evidently this is absolute nonsense, a given integer n is either prime or it isn't, but this will turn out to be useful when suitably formulated[6], so let's let things be for now. Indeed it was this statement of Gauss that led, as we shall see, to Cramér's probabilistic approach for understanding the distribution of prime numbers, which underpins most of the heuristic reasoning still used in the subject today. So the 'expected' number of primes up to x is

$$\approx \sum_{2\le n\le x} \frac{1}{\log n} = \int_2^x \frac{\mathrm{d}t}{\log t} + O(1); \tag{5}$$

we denote this integral as $\mathrm{Li}(x)$. Integrating by parts gives

$$\frac{x}{\log x} + \frac{x}{\log^2 x} + O\left(\frac{x}{\log^3 x}\right); \tag{6}$$

a somewhat different expected answer from that given by the heuristic based on the Sieve of Eratosthenes (see (2)). So which one is correct? The argument that gives (2), based on the assumption that the error terms do not 'accumulate' (that is that there is a sufficient amount of 'independence' in sieving by different primes)? Or Gauss's 'probability' statement, based on viewing primes as having a specific 'density'?

In 1851, Chebychev proved that if $\pi(x)/(x/\log x)$ does tend to a limit as x goes to infinity, then that limit must be one[7]. However he was unable to show that the limit exists! Moreover, we can see in the following table that recent computational evidence agrees well with Gauss's prediction:

[4]It was Wiles' work on this conjecture that led to his recent spectacular attack on Fermat's Last Theorem.

[5]The 'as a boy' preface by Gauss is probably more fact than arrogance: Indeed Gauss's great classic "*Disquisitiones Arithmeticae*" was completed by the time he was in his mid-twenties!

[6]as well as somewhat less ridiculous!

[7]Thus if either of predictions above is correct, it must be Gauss's.

x	$\pi(x)$	$[\mathrm{Li}(x) - \pi(x)]$
10^8	5761455	754
10^9	50847534	1701
10^{10}	455052511	3104
10^{11}	4118054813	11588
10^{12}	37607912018	38263
10^{13}	346065536839	108971
10^{14}	3204941750802	314890
10^{15}	29844570422669	1052619
10^{16}	279238341033925	3214632
4×10^{16}	1075292778753150	5538861

The number of primes, $\pi(x)$, up to x.

Riemann, the (immediate) successor to Gauss and Dirichlet's chair in Berlin, 'resolved' this dispute with his extraordinary eight page memoir, presented to the Berlin Academy in 1859. In it, Riemann proposed a careful study of the function

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} \qquad (\mathrm{Re}(s) > 1)$$

considered now as a function of the complex variable s, analytically continued to the whole-plane, though with a simple pole at $s = 1$. From the formula

$$\sum_{\substack{p \text{ prime} \\ m \geq 1}} \frac{\log p}{p^{ms}} = -\frac{\zeta'(s)}{\zeta(s)} \qquad \text{for } \mathrm{Re}(s) > 1 \tag{7}$$

(which one can derive from (3)), one can use the discontinuous integral

$$\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} y^s \frac{\mathrm{d}s}{s} = \begin{cases} 0 & \text{if } 0 < y < 1, \\ \frac{1}{2} & \text{if } y = 1\ , \\ 1 & \text{if } y > 1\ , \end{cases}$$

to pick out the terms with $p^m \leq x$ in (7)[8], and deduce that if x is not a power of a prime then

$$\sum_{p^m \leq x} \log p = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, \mathrm{d}s \qquad \text{for } c > 1.$$

By moving the vertical line of integration away to infinity on the left, this can be expressed as a sum of residues of $-(\zeta'/\zeta)(s)x^s/s$ at its poles. In fact, since the pole of $\zeta(s)$ at $s = 1$ contributes x, we arrive at the formula, for $x \geq 2$,

$$\sum_{p^m \leq x} \log p = x - \sum_{\zeta(\rho)=0} \frac{x^\rho}{\rho} + O(1). \tag{8}$$

If we assume that all of the zeros satisfy $\mathrm{Re}(\rho) < 1$, and that they are not 'too dense' in any part of the complex plane (in particular near to $\mathrm{Re}(s) = 1$) then (8)

[8] by taking $y = x/p^m$

leads to the asymptotic formula

$$\sum_{p^m \le x} \log p \sim x \quad (\text{as } x \to \infty). \tag{9}$$

By partial summation, this confirms Gauss's prediction, (6).

It was not until 1896 that Hadamard and de la Vallée Poussin filled in the details of the outline above, to prove (9), the "Prime Number Theorem". Actually Riemann went much further with his extraordinary approach, conjecturing what we now call

The Riemann Hypothesis: If $\zeta(s) = 0$ with $0 \le \mathrm{Re}(s) \le 1$ then $\mathrm{Re}(s) = 1/2$.

Using (8) one can deduce from the Riemann Hypothesis (RH) that

$$\sum_{p^m \le x} \log p = x + O(\sqrt{x} \log^2 x),$$

which implies, via partial summation,

$$\pi(x) = \int_2^t \frac{\mathrm{d}t}{\log t} + O(\sqrt{x} \log x) \tag{10}$$

This extraordinary and profound connection between the count of the number of primes and the zeros of a complex analytic function energized thinking in mathematics at the beginning of this century.[9] As Ingham put it in 1932,

> "*The solution*[10] *... may be held to be unsatisfactory in that it introduces ideas very remote from the original problem, and it is natural to ask for a proof of the prime number theorem not depending on the theory of functions of a complex variable ... It seems unlikely that a genuinely 'real variable' proof will be discovered.*"

Ingham goes on to note how, via the formula (8), the prime number theorem is, essentially, *equivalent* to the fact $\zeta(\rho) = 0 \Rightarrow \mathrm{Re}(\rho) < 1$, and so any proof **must** use some complex analysis.

How wrong Ingham was (as well as Hardy, Bohr, and many others)! In 1949 Selberg and Erdös showed that it **is** possible to give an 'elementary proof' of the prime number theorem[11].

Since his student days Cramér had been interested in the size of gaps between consecutive primes. From (10) one can easily deduce that

$$p_{n+1} - p_n = O(\sqrt{p_n} \log^2 p_n)$$

where $p_1, p_2, \ldots$ is the sequence of prime numbers. In 1920 Cramér sharpened this to

$$p_{n+1} - p_n = O(\sqrt{p_n} \log p_n)$$

[9] Indeed RH and subsequent investigations comprise the eighth of Hilbert's 23 problems announced as a challenge for the forthcoming century at the International Congress of Mathematicians held in Paris in 1900.

[10] that is, to the problem of proving the prime number theorem (9)

[11] For a thorough analysis of how and why one can avoid complex analysis, and indeed any type of 'infinite summation' in such a proof see Ingham's elegant Math Review [9].

On 25th October 1920 G. H. Hardy read Cramér's paper "*On the distribution of primes*" to the Cambridge Philosophical Society. Here Cramér develops a 'statistical approach' to this question showing that for any fixed $\varepsilon > 0$

$$p_{n+1} - p_n = O(p_n^{\varepsilon})$$

for 'most' p_n: in fact for all but at most $x^{1-3\varepsilon/2}$ of the primes $p_n \le x$.

By the mid 1930's, Cramér had shown that

$$p_{n+1} - p_n = o(\log^3 p_n)$$

for all but at most $o(x/\log^4 x)$ primes $p_n \le x$; and further that

$$\sum_{\substack{p_n \le x \\ p_{n+1}-p_n \ge y}} (p_{n+1} - p_n) = O\left(\frac{x \log^3 x}{y \log y}\right) \tag{11}$$

This kind of statistical result brought a whole new dimension to these considerations. Evidently it had to be important to understand the 'usual' behaviour of primes as well as the extreme cases. This was a big first step by Cramér and, within a few years, Selberg significantly developed this type of idea, bringing this kind of approach to maturity.

Various upper bounds on gaps between primes have been proved without the assumption of an unproved hypothesis. In 1930 Hoheisel proved that $p_{n+1} - p_n = O(p_n^{1-\delta})$ for some constant $\delta > 0$. In 1936 Tchudakoff, and in 1937 Cramér, showed one could take $\delta = 1/4 - \varepsilon$. Successively bigger values of δ have been given since, the latest results of R. C. Baker and Harman (1994) being close to what can be proved assuming RH: they prove, unconditionally,

$$p_{n+1} - p_n = O\left(p_n^{\frac{1}{2}+\frac{7}{200}}\right).$$

Mathematicians have also attempted to show that there must be large gaps between consecutive primes. One way to do this is to find a long run of consecutive integers which each have a small prime factor[12]. For example, $n! - j$ is divisible by j for $2 \le j \le n$ and this gives rise to a gap of length $\ge n$ between consecutive primes[13]. One can do a little better from (9), which implies that there must be gaps $> \{1 + o(1)\} \log x$ between some pair of consecutive primes $\le x$. In 1931 Westzynthius improved this to

$$\limsup_{n \to \infty} \ (p_{n+1} - p_n)/\log p_n = \infty.$$

Following work of Erdös and Rankin we now know that there is an infinite sequence of primes p_n for which

$$p_{n+1} - p_n > \frac{c(\log p_n)(\log\log p_n)(\log\log\log\log p_n)}{(\log\log\log p_n)^2}$$

for some fixed constant $c > 0$[14].

It is also interesting to try to determine the minimal order of gaps between consecutive primes. A very famous open question in number theory is whether there

[12] Indeed most papers concerning this question have used this simple idea

[13] Thus, if $x = n!$ we have an interval of length $\ge n \sim \log x/\log\log x$ (by Stirling's formula) between consecutive primes near to x

[14] In the intervening half century, the constant c in this result has been repeatedly improved but not the function itself. Erdös offers \$5000 for such an improvement!

are infinitely many 'twin primes', that is prime pairs $p, p+2$. In 1920 the Norwegian mathematician Viggo Brun showed that they are nowhere near as numerous as the primes; in that

$$\sum_{\substack{p,p+2\\ \text{both primes}}} \frac{1}{p} \quad \text{converges, whereas} \qquad \sum_{p \text{ prime}} \frac{1}{p} \quad \text{diverges.}$$

In 1923, Hardy and Littlewood found a new perspective from which to study such questions and conjectured[15]

$$\#\{p \le x : p,\ p+2 \text{ both prime}\} \sim \frac{1}{2} \prod_{p \ge 3} \left(1 + \frac{1}{p(p-1)}\right) \cdot \frac{x}{\log^2 x}. \tag{12}$$

More generally, for any k-tuple of distinct integers $a_1, a_2, \dots, a_k$ they conjectured that

$$\#\{p \le x : p + a_1, p + a_2, \dots, p + a_k \text{ are all prime}\} \sim C(\mathbf{a}) \cdot \frac{\mathbf{x}}{\log^{\mathbf{k}} \mathbf{x}} \tag{13}$$

where the constant $C(\mathbf{a})$ depends only on $a_1, a_2, \dots, a_k$. One has to be a little careful here — evidently p and $p+1$ are only simultaneously prime if $p = 2$, since one of the two numbers must be even. Thus we only make the above "prime k-tuplets conjecture" when $a_1, a_2, \dots, a_k$ is an 'admissible set'; that is that there is no such obstruction $\mod 2$ or $\mod 3$ or mod any prime.

Actually no-one has yet proved that there are infinitely many primes p_n such that $p_{n+1} - p_n < \frac{1}{10} \log p_n$, a question that seems to be much more difficult than it looks.

In 1937 Cramér decided to try to guess at the true order of $\max_{p_n \le x}(p_{n+1} - p_n)$ using a sophisticated heuristic argument based on Gauss's observations above, but quite unlike anything seen before in analytic number theory. In Cramér's words:

> "In investigations concerning the asymptotic properties of arithmetic functions, it is often possible to make an interesting heuristic use of probability arguments. If, *e.g.*, we are interested in the distribution of a given sequence S of integers, we then consider S as a member of an infinite class C of sequences, which may be concretely interpreted as the possible realizations of some game of chance. It is then in many cases possible to prove that, *with a probability* $= 1$, a certain relation R holds in C, *i.e.* that in a definite mathematical sense "almost all" sequences of C satisfy R. Of course we cannot in general conclude that R holds for the particular sequence S, but results suggested in this way may sometimes afterwards be rigorously proved by other methods.
>
> With respect to the ordinary prime numbers, it is well known that, roughly speaking, we may say that the chance that a given integer n should be a prime is approximately $\frac{1}{\log n}$. This suggests that by considering the following series of independent trials we should obtain sequences of integers presenting a certain analogy with the sequence of ordinary prime numbers p_n.

[15] In 1871 Sylvester had conjectured a similar asymptotic formula but with a slightly different constant. Hardy and Littlewood contrast their conjecture with Sylvester's in the same way that we have contrasted (6) with (2). They explained such discrepancies between different heuristic models by stating, "*Probability is not a notion of pure mathematics, but of philosophy or physics.*"

Let $U_1, U_2, U_3, \ldots$ be an infinite series of urns containing black and white balls, the chance of drawing a white ball from U_n being $\frac{1}{\log n}$ for $n > 2$, while the composition of U_1 and U_2 may be arbitrarily chosen. We now assume that one ball is drawn from each urn, so that an infinite series of alternately black and white balls is obtained. If P_n denotes the number of the urn from which the nth white ball in the series was drawn, the numbers $P_1, P_2, \ldots$ will form an increasing sequence of integers, and we shall consider the class C of all possible sequences (P_n). Obviously the sequence S of ordinary prime numbers (p_n) belongs to this class.

We shall denote by $\Pi(x)$ the number of those P_n which are $\leq x$, thus forming an analogy to the ordinary notation $\pi(x)$ for the number of primes $p_n \leq x$. Then $\Pi(x)$ is a random variable, and if we denote by z_n a variable taking the value 1 if the nth urn gives a white ball and the value 0 in the opposite case, we have

$$\Pi(x) = \sum_{n \leq x} z_n,$$

and it is easily seen that the mean value of $\Pi(x)$ is, for large values of x, asymptotically equal to $\mathrm{Li}(x)$. It is, however, possible to obtain much more precise information concerning the behaviour of $\Pi(x)$ for large values of x. As a matter of fact, it may be shown that, *with a probability* $= 1$, the relation

$$\limsup_{x \to \infty} \frac{|\Pi(x) - \mathrm{Li}(x)|}{\sqrt{2x} \cdot \sqrt{\frac{\log \log x}{\log x}}} = 1$$

is satisfied. With respect to the corresponding difference $\pi(x) - \mathrm{Li}(x)$ in the prime number problem, it is known that, if the Riemann hypothesis is assumed, the true maximum order of this difference lies between the functions $\frac{\sqrt{x}}{\log x}$ and $\sqrt{x} \cdot \log x$. It is interesting to find that the order of the function occurring in the denominator in the above equation falls inside this interval of indetermination.

We shall now consider the order of magnitude of the difference $P_{n+1} - P_n$. Let $c > 0$ be a given constant and let E_m denote the event that black balls are obtained from all urns U_{m+v} with $1 \leq v \leq c(\log m)^2$. Then it is seen that the following two events have the same probability:
a) The inequality

$$P_{n+1} - P_n > c(\log P_n)^2 \qquad (*)$$

is satisfied for an infinity of values of n, and
b) An infinite number of the events E_m are realized."

Cramér next proves that the probability of the event E_m occurring is $\asymp m^{-c}$. Citing Cantelli he continues,

"Thus the probability of an infinite number of solutions of the inequality (*) is equal to zero if $c > 1$ and to one if $c < 1$. Combining these two results, we obtain the following theorem: *With a probability* $= 1$, *the relation*

$$\limsup_{n \to \infty} \frac{P_{n+1} - P_n}{(\log P_n)^2} = 1$$

is satisfied. — Obviously we may take this as a suggestion that, for the particular sequence of ordinary prime numbers p_n, some similar relation may hold."

So what Cramér seems to be suggesting, on probabilistic grounds, is that the largest gap between consecutive primes $\le x$ is $\sim \log^2 x$; more precisely,

$$\max_{p_n \le x} (p_{n+1} - p_n) \sim \log^2 x. \tag{14}$$

This statement (or the weaker $O(\log^2 x)$) is known as '*Cramér's Conjecture.*' Shanks reformulated this statement to suggest that the first occurence of a gap between consecutive primes of size $> g$ would occur with $p_n = \mathrm{e}^{\{1+o(1)\}\sqrt{g}}$. Computations of gaps between consecutive primes indicate that Cramér may well have been correct:

p_n	$p_{n+1} - p_n$	$(p_{n+1} - p_n)/\log^2 p_n$
31397	72	.6715
370261	112	.6812
2010733	148	.7025
20831323	210	.7394
25056082087	456	.7953
2614941710599	652	.7975
19581334192423	778	.8177

Record-breaking gaps between primes, up to 10^{14}

Gauss's assertion is really about the number of primes in a short interval near to a given value of x: that is

$$\pi(x+y) - \pi(x)$$

where y is 'small' compared to x.

Cramér's analysis of $\Pi(x)$ provided 'expected results' which are similar to those obtained from assuming the Riemann Hypothesis, for y about the same size as x. Since his predictions fit the facts so well for y around size x, we now consider analogous arguments for smaller values of y.

The independent random variables z_j, with j close to x, satisfy

$$\mathrm{Prob}(z_j = 1) \approx p \quad \text{and} \quad \mathrm{Prob}(z_j = 0) \approx 1 - p$$

for each j, where $p = 1/\log x$. It is well known that if, for such binomial distributions, we let $n \to \infty$ while keeping np fixed, then we get a Poisson distribution for the sum $z_1 + z_2 + \cdots + z_n$. Thus for any fixed $\lambda > 0$ and integer $k \ge 0$, we have

$$\#\{\text{integers } x \le X : \ \Pi(x + \lambda \log x) - \Pi(x) = k\} \sim \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} X \tag{15}$$

as $X \to \infty$, with probability 1 in C.

In 1966 Gallagher investigated some consequences of Hardy and Littlewood's conjecture (13) about the distribution of prime k-tuplets. By assuming (13) holds in an appropriate 'uniform' way, Gallagher deduced that (15) holds for π, the sequence of primes, as well as 'almost certainly' for Π. Yet another success for predictions arising out of Cramér's model, and especially interesting because the ideas motivating Hardy and Littlewood's conjecture are quite different from the ideas motivating the Riemann Hypothesis.

Thus, for y large ($\asymp x$), Cramér's model confirms what is known from assuming the Riemann Hypothesis, and for y small ($\asymp \log x$), it confirms what is known from assuming a uniform version of the Hardy-Littlewood conjecture.

It is not difficult to show that if $y/\log^2 x \to \infty$ then

$$\Pi(x+y) - \Pi(x) \ = \int_x^{x+y} \frac{\mathrm{d}t}{\log t} + O(\sqrt{y}), \tag{16}$$

and thus

$$\Pi(x+y) - \Pi(x) \ = \ \{1+o(1)\}\frac{y}{\log x}, \tag{17}$$

holds with probability 1. In 1943 Selberg showed that if $y/\log^2 x \to \infty$ then

$$\pi(x+y) - \pi(x) \ = \ \{1+o(1)\}\frac{y}{\log x} \tag{18}$$

for all but $o(x)$ integers $x \le X$.[16] So, in this intermediate range of y-values, Cramér's model yet again predicts what we believe to be true for other (good) reasons.

With so much evidence to support predictions that come from Cramér's model it came as a great surprise when, in 1985, Maier proved a result that actually contradicts what one expects from Cramér's model:

As a consequence of (17) one can deduce that, for any fixed $N > 2$,

$$\Pi(x + \log^N x) - \Pi(x) \ \sim \ \log^{N-1} x \tag{19}$$

as $x \to \infty$, with probability 1. Cramér's 'philosophy' leads us to expect that (19) holds for the sequence of primes. However Maier proved that this is not true. Specifically, that there exists a constant $\delta_N > 0$ such that

$$\pi(x_+ + \log^N x_+) - \pi(x_+) \ > \ (1+\delta_N)\log^{N-1} x_+$$

for arbitrarily large values of x_+, and such that

$$\pi(x_- + \log^N x_-) - \pi(x_-) \ < \ (1-\delta_N)\log^{N-1} x_-$$

for arbitrarily large values of x_-.[17]

Maier's result is totally unexpected from the perspective of Cramér's original model. What he does is to brilliantly combine the approaches of Eratosthenes and of Gauss to exploit that old inconsistency between (6) and (2). To try to guess at the number of primes between x and $x+y$, Maier first removes those integers that have a small prime factor (*following Eratosthenes*), and only then does he apply density arguments (*following Gauss*). This surely must provide a more accurate prediction than the previous model — what is surprising is that, on occasion, it provides a quite different prediction ...

[16] In the same paper Selberg also improved the upper bound in (11) to $O((x/y)\log^2 x)$.

[17] The values of $x_\pm$ given by Maier's proof are scarce, so this is still consistent with Selberg's result (18).

Let $Z_3, Z_4, \dots$ be a sequence of independent random variables with $Z_n = 0$ whenever n has a prime factor $\le T$. When n is free of prime factors $\le T$ then

$$\text{Prob}(Z_n = 1) = \prod_{p \le T} \left(\frac{p}{p-1}\right) \cdot \frac{1}{\log n}$$

$$\text{Prob}(Z_n = 0) = 1 - \prod_{p \le T} \left(\frac{p}{p-1}\right) \cdot \frac{1}{\log n},$$

where T is a parameter to be chosen appropriately. Note that when $T = 1$ this is exactly Cramér's model; however, we shall take T to be at least some power of $\log x$. This new model has several important advantages over Cramér's. For example, from Cramér's model one expects that there are $\sim x/\log^2 x$ prime pairs $p, p+1$ up to x, whereas our new model recognizes that one of these two numbers must be even[18]. Also Cramér's model leads one to expect that there are $\sim x/\log^2 x$ twin primes $p, p+2$ up to x, whereas our new model will lead us to predict (12), which is presumably correct.

Anyway, from our new model we believe that there are

$$\approx \frac{1}{\log x} \prod_{p \le T} \left(1 - \frac{1}{p}\right)^{-1} \sum_{\substack{x < n \le x+y \\ p \nmid n \text{ for all } p \le T}} 1$$

primes in $(x, x+y]$; and that this should be an asymptotic estimate if $y/\log^2 x \to \infty$. If this is consistent with the prediction (17) arising from Cramér's model then we must have

$$\sum_{\substack{x \le n \le x+y \\ p \nmid n \text{ for all } p \le T}} 1 \sim \prod_{p \le T} \left(1 - \frac{1}{p}\right) \cdot y. \tag{20}$$

However, if we take $T = y^{1/2+o(1)}$, and choose x so that it is divisible by $\prod_{p \le T} p$ then the left hand side of (20) equals[19]

$$\sum_{\substack{n \le y \\ p \nmid n \text{ for all } p \le T}} 1 \approx \pi(y) \sim \frac{y}{\log y},$$

whereas the right side of (20) is

$$\prod_{p \le T} \left(1 - \frac{1}{p}\right) \cdot y \sim \frac{y \mathrm{e}^{-\gamma}}{\frac{1}{2}\log y} \sim \frac{2\mathrm{e}^{-\gamma} y}{\log y}.$$

So we see that old inconsistency between (6) and (2) appearing again!

In his very ingenious paper, Maier was able to exploit this inconsistency to prove the existence of arbitrarily large x_+ and x_- above, a severe blow to Cramér's model. Moreover, with our new model above, Cramér's arguments suggest that

$$\max_{p_n \le x} (p_{n+1} - p_n) \gtrsim 2\mathrm{e}^{-\gamma} \log^2 x,$$

which contradicts Cramér's conjecture (14)! The computational evidence alone (see above) would not lead one to predict that (14) errs on the small side, but the

[18] however experienced researchers have long known not to apply Cramér's model to try to understand inappropriate problems!

[19] since prime $p \le T$ divides $x + k$ if and only if it divides k

data collected so far is very limited, and there are now a number of independent computors trying to find examples with $(p_{n+1}-p_n)/\log^2 p_n > 1$.

Let $\pi(x;q,a)$ denote the number of primes $\le x$, that belong to the arithmetic progression $a \pmod q$. Euler's 'totient function' $\phi(q)$ is defined to be the number of integers $a, 1 \le a \le q$ for which $(a,q)=1$. We might expect the primes up to x to be distributed equally amongst the $\phi(q)$ arithmetic progressions $a \pmod q$ with $(a,q)=1$[20], so that

$$\pi(x;q,a) \sim \frac{\pi(x)}{\phi(q)} \tag{21}$$

whenever a and q are fixed integers for which $(a,q)=1$, as $x \to \infty$. This was proved by de la Vallée Poussin in 1896, combining ideas used in the proof of the Prime Number Theorem with ideas of Dirichlet.

Riemann's analytic approach may be modified to investigate the distribution of primes in arithmetic progressions. A suitable generalization of Riemann's Hypothesis implies that (21) holds uniformly for

$$2 \le q \le x^{\frac{1}{2}}/\log^{2+\varepsilon} x \quad with \quad (a,q)=1\ , \tag{22}$$

for any fixed $\varepsilon > 0$. A result of this strength would allow us to answer many important questions of number theory and so one wishes to establish an estimate like (21) for each $(a,q)=1$ with values of q and x in as wide a range as possible.

The best such result was proved by Walfisz in 1936, applying an idea of Siegel. He showed that (21) holds uniformly for any $q \le \log^N x$, for any fixed $N > 0$. In the mid-sixties Bombieri[21] used the 'large sieve' to show that (21) holds uniformly for 'almost all' moduli q in the range (22) for all $(a,q)=1$, for some fixed $\varepsilon > 0$[22].

One can analyze the distribution of primes in arithmetic progressions using a probabilistic model much as Cramér did for primes in intervals: The total number of primes up to x, summing over all arithmetic progressions $a \pmod q$ with $(a,q)=1$, is $\pi(x)$ minus the number of distinct prime factors of q, which is $\sim x/\log x$. The total number of integers up to x in those arithmetic progressions is $\sim \phi(q)x/q$, and so the probability that one such integer, chosen at random, is prime is $\sim \phi(q)/(q\log x)$. Setting up our probability space as before we expect that the estimate

$$\pi(x;q,a) = \frac{\pi(x)}{\phi(q)} + O\left(\left(\frac{x}{q}\right)^{\frac{1}{2}}\log(qx)\right), \tag{23}$$

and thus (21), holds uniformly in the range

$$2 \le q \le x/\log^{2+\varepsilon} x \quad with \quad (a,q)=1 \tag{24}$$

for any fixed $\varepsilon > 0$. However, one can modify the method of Maier to show that (23) cannot hold in at least part of this range.

With hindsight it is rather ambitious to expect (21) to hold for x as small as, say, $q\log^3 q$, for all $(a,q)=1$. On the other hand it is known that (21) holds for 'almost all' pairs (a,q) in the range (24), and it was widely believed that (21) held for 'almost all' q and all $(a,q)=1$ in the range (24)[23]. However, Friedlander and I recently

[20] there is no more than one such prime if $(a,q)>1$ since then (a,q) divides every number in the arithmetic progression,

[21] improving on work of Linnik, Renyi, Roth, Vinogradov and others

[22] so more-or-less proving what follows from the generalized Riemann Hypothesis, for 'most' moduli q

[23] in other words, that the range in Bombieri's result can be extended from (22) to (24).

showed that even this averaged conjecture is untrue, by applying an appropriate modification of Maier's ideas. Specifically, in direct analogy with Maier's result for $\pi(x+\log^N x)-\pi(x)$, we showed that for any fixed $N>0$ there exists a constant $\delta_N>0$ such that for any modulus q with 'not too many' small prime factors there exist arithmetic progressions $a_\pm \pmod q$ and values $x_\pm \in [\phi(q)\log^N q, 2\phi(q)\log^N q]$ such that

$$\pi(x_+;q,a_+) \ > \ (1+\delta_N)\frac{\pi(x_+)}{\phi(q)}$$

and

$$\pi(x_-;q,a_-) \ < \ (1-\delta_N)\frac{\pi(x_-)}{\phi(q)} \ .$$

Thus, for any modulus q with 'few' small prime factors, there exist values of x around $\phi(q)\log^N q$ and of a with $(a,q)=1$, for which (21) fails.

In 1992, Friedlander and I went somewhat further. First we showed that (21) fails more often, the more prime factors that a has[24]. However we showed that even when a has very few prime factors, (21) can still fail. For example, the primes $\equiv 1 \pmod q$ do not satisfy (21) uniformly in the range $q<x/\log^N x$, for any fixed $N>0$. It also seems that Maier's idea may effect our understanding of the prime k-tuplets conjecture (13), and of the distribution of the zeros of $\zeta(s)$ on the line $\mathrm{Re}(s)=1/2$ (assuming RH).

It does appear that we need to re-appraise just about every conjecture concerning the distribution of prime numbers, in the light of Maier's revolutionary idea. Presumably we will remain unable to fully understand the finer details until a model is proposed that adequately accounts for both the sieve of Eratosthenes, and Gauss's density statement. But perhaps it is hopeless, perhaps Euler was correct when he wrote

> *"Mathematicians have tried in vain to discover some order in the sequence of prime numbers but we have every reason to believe that there are some mysteries which the human mind will never penetrate."*
>
> — L. Euler (1770).

To conclude, just in case you are starting to think that everything proved recently concerning the distribution of prime numbers implies that what has long been believed is wrong, I'd like to tell you about a beautiful recent result of Balog:

Balog has shown that, in an appropriate sense, the prime k-tuplets conjecture (13) holds on average (in the tradition of Cramér). One delightful consequence is the existence of infinitely many squares and triangles of primes in 3-term arithmetic progressions as in this picture:

Balog even shows that there are infinitely many such n-dimensional cubes and tetrahedrons.

A more detailed survey of these ideas, together with more applications and the fullest strength of the results that have been proved, will appear in the author's

[24] something that had not shown up in previous considerations.

forthcoming article for the Proceedings of the 1994 International Congress of Mathematicians.

ACKNOWLEDGEMENTS: I'd like to thank Red Alford, Ken Ono and Carl Pomerance for their comments on earlier drafts of this article.

RECOMMENDED FURTHER READING

Two essays that breathe life into the beauty and mystery of the distribution of prime numbers are "*Prime Territory*" by E. Bombieri which appeared in The Sciences, Sept/Oct 1992, 30-36; and "*The first 50 million prime numbers*" by D. Zagier in Math. Intell. 1977, 7-19. P. Ribenboim's "*Book of Prime Number Records*"[25] (Springer-Verlag), glows with the fun of learning about primes, though many details are suspended.

There are two elegant treatments of the classical theory: H. Davenport's "*Multiplicative Number Theory*" (Springer-Verlag), and A. E. Ingham's "*Distribution of Prime Numbers*" (C.U. Press) which is older but still worth looking at for its eloquence.

For more about the Riemann zeta-function, the classic "*The Theory of the Riemann Zeta-function*" (O.U. Press) by E. C. Titchmarsh is essential reading. To learn about sieving one cannot do better than study the relevant material in A. Selberg's "*Collected Works*", some of which has not appeared elsewhere. E. Bombieri's "*Le Grand Crible en la theorie analytique des nombres*" (Astérisque, 18) clearly describes the large sieve with lots of applications, and H. L. Montgomery's "*Multiplicative Number Theory*" (Springer Lecture Notes, 227) is indispensable for the serious researcher.

REFERENCES

[1] Balog, A. 1990 The prime k-tuplets conjecture on average. Analytic Number Theory (ed. B.C. Berndt, H.G. Diamond, H. Halberstam, A. Hildebrand), (Birkhäuser, Boston), 165–204.
[2] Cramér, H. 1936 On the order of magnitude of the difference between consecutive prime numbers. *Acta Arith.* **2**, 23–46.
[3] Cramér, H. 1935 Prime numbers and probability. *Skand. Mat.-Kongr.* **8**, 107-115.
[4] Cramér, H. 1920 On the distribution of primes. *Proc. Camb. Phil. Soc.* **20**, 272-280.
[5] Cramér, H. 1920 Some theorems concerning prime numbers. *Arkiv för Mat. Astr. o. Fys.* **15**, #5, 1-32.
[6] Friedlander, J. and Granville, A. 1989 Limitations to the equi–distribution of primes I. *Ann. Math.* **129**, 363–382.
[7] Friedlander, J. and Granville, A. 1992 Limitations to the equi–distribution of primes III. *Comp. Math.* **81**, 19–32.
[8] Friedlander, J. and Granville, A. 1992 Relevance of the residue class to the abundance of primes. *Proceedings of the Amalfi Conference*, 95–104.
[9] Hardy, G. H. and Littlewood, J. E. 1923 Some problems on partitio numerorum III. On the expression of a number as a sum of primes. *Acta Math.* **44**, 1–70.
[10] Ingham, A. E. 1949, Mathematical Reviews **10**, 595 b & c.
[11] Maier, H. 1985 Primes in short intervals. *Michigan Math. J.* **32**, 221–225.
[12] Selberg, A. 1943 On the normal density of primes in small intervals and the difference between consecutive primes. *Arch. Math. Naturvid.* **47** 87–105.

[25] and the abridged paperback version, "*The Little Book of Big Primes*"

Andrew Granville, Department of Mathematics, University of Georgia, Athens, GA 30602, USA

E-mail address: andrew@math.uga.edu