Twin primes (seem to be) more random than primes

Richard P. Brent

Australian National University
and University of Newcastle

25 October 2014

Abstract

Cramér's probabilistic model (1936) gives a useful way of predicting the distribution of primes. However, the model is not perfect, as shown by Cramér himself (1920), Maier (1985), and Pintz (2007).

With appropriate modifications (Hardy and Littlewood), Cramér's model can also be used for twin primes. In fact, computational evidence suggests that the model is better for twin primes than for primes. In this sense, twin primes appear to behave more randomly than primes. This can be explained by the connection between the Riemann zeta function and the primes - the zeros of the zeta function constrain the behaviour of the primes, but there does not appear to be any analogous constraint on the behaviour of twin primes.

In the talk I will outline some of the results mentioned above, and present some numerical evidence for the claim that twin primes are more random than primes.

Cramér's model

> *"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 $1/\log x$."*
>
> *Letter from Gauss to Encke, 1849*

In the 1930s, Harald Cramér introduced the following "model" of the primes (as paraphrased by Soundararajan):

> ***CM**: The primes behave like independent random variables $X(n)$ $(n \geq 3)$ with $X(n) = 1$ (the number n is "prime") with probability $1/\log n$, and $X(n) = 0$ (the number n is "composite") with probability $1 - 1/\log n$.*
>
> *Cramér (1936)/Soundararajan (2006)*

Notation: in this talk, $\log x = \ln x$, and $\log_2 x = \ln\ln x$, $\log_3 x = \ln\ln\ln x$, etc.

Local limitations of Cramér's model

Of course, Cramér did not intend his model to be taken literally. "Locally", the primes do not behave like independent random variables. For example,

$$X(n) = 1 \implies X(n+1) = 0 \;\text{ for }\; n \geq 3.$$

The frequencies of k-tuples of primes $\leq N$ are expected to be $\sim CN/(\log N)^k$, where C is a constant that can be computed from the given k-tuple [Hardy and Littlewood, 1923]. For example, the frequency of twin primes (primes p such that $p+2$ is also prime) is expected to be

$$\pi_2(N) \sim \frac{2c_2 N}{(\log N)^2},$$

where $2c_2 > 1$ (the constant 1 is implied by Cramér's model). More precisely,

$$2c_2 = 2 \prod_{p \geq 3 \text{ prime}} \left(1 - \frac{1}{(p-1)^2}\right) \approx 1.3203236.$$

Semi-local behaviour of $\pi(x)$

Let $\pi(x)$ be the prime-counting function. For which functions $\Psi(x) \ll x$ is it true that

$$\pi(x + \Psi(x)) - \pi(x) \sim \frac{\Psi(x)}{\log x} \text{ as } x \to \infty \text{ ?} \qquad (1)$$

To paraphrase, how large does an interval need to be before the primes can be guaranteed to have their "expected" distribution asymptotically?

Huxley (1972) and Heath-Brown (1980) showed that one can take $\Psi(x) = x^{7/12}$. The Riemann Hypothesis (RH) implies that one can $\Psi(x) = x^{1/2+\varepsilon}$.

Huxley (1972) showed that (1) is true *for almost all x* if $\Psi(x) = x^{1/6+\varepsilon}$. Selberg (1943), assuming RH, showed this for $\Psi(x)/(\log x)^2 \to \infty$. The same result follows from Cramer's model CM with probability 1.

Semi-local behaviour – Rankin's theorem

In the other direction, Rankin (1938) showed that (1) is false if

$$\Psi(x) = c(\log x)(\log_2 x)(\log_4 x)(\log_3 x)^{-2},$$

for a sufficiently small $c > 0$. This improved on a sequence of earlier results by Erdős and others.

Maynard (2014) and Ford, Green, Konyagin and Tao (2014) recently improved on Rankin just so slightly. They showed that

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{(\log p_n)(\log_2 p_n)(\log_4 p_n)(\log_3 p_n)^{-2}} = \infty.$$

There is a big gap between the upper and lower bounds on $\Psi(x)$, even assuming RH. Ignoring epsilons and log log factors, the upper bounds are close to $x^{1/2}$, while the lower bounds are close to $\log x$.

Maier's theorem

It was a great surprise[1] when Maier (1985) showed that (1) is false for $\Psi(x) = (\log x)^\lambda$ and any $\lambda > 1$. In particular,

$$\limsup_{x\to\infty} \frac{\pi(x+\Psi(x)) - \pi(x)}{\Psi(x)/\log x} > 1$$

and

$$\liminf_{x\to\infty} \frac{\pi(x+\Psi(x)) - \pi(x)}{\Psi(x)/\log x} < 1.$$

Maier's theorem contradicts CM if $\lambda > 2$.

It does not contradict Selberg's result, since the exceptional values of x are rare.

[1] According to Carl Pomerance (2009), "both Erdős and Selberg were astonished".

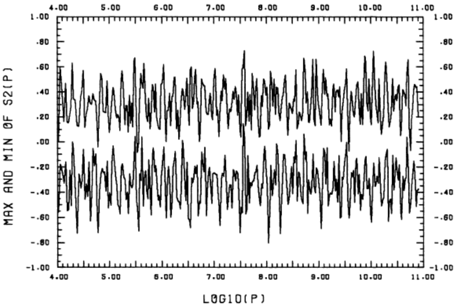
Observations on primes

In 1975 I computed the error in various approximations to the counts of primes and of twin primes $\le x$.

For the primes I considered Riemann's approximation

$$R(x) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \mathrm{Li}(x^{1/k}) = 1 + \sum_{k=1}^{\infty} \frac{(\log x)^k}{k!k\zeta(k+1)} .$$

FIGURE 1
RIEMANN'S APPROXIMATION

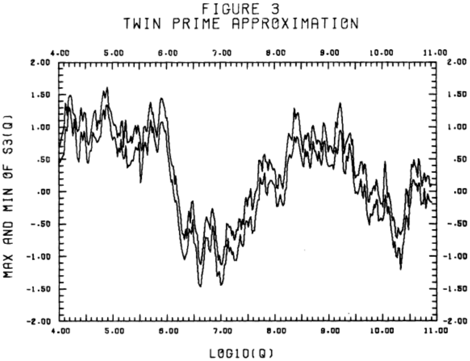


The same for twin primes

Here is the corresponding graph for an approximation

$$L_2(x) = 2c_2 \int_2^x \frac{dt}{\log^2 t}$$

to the twin-prime counting function $\pi_2(x)$.



Primes vs twin primes

You will notice that the two graphs look quite different.

For primes the error oscillates very fast and there are many zero-crossings (in the graph I only plotted the upper and lower envelopes). It does not look like a Bernoulli random walk.

For twin primes there is a slow oscillation that looks much more like what one expects for a Bernoulli random walk.

A global limitation of CM

To put our observations about primes on a sound theoretical footing, we could consider the mean square errors

$$\frac{1}{X}\int_3^X (\pi(x) - R(x))^2\,dx \quad \text{or} \quad \frac{1}{X}\int_3^X (\pi(x) - \mathrm{Li}(x))^2\,dx \tag{2}$$

and compare these with the predictions from CM.
It is simpler to consider CM's prediction for

$$\frac{1}{X}\int_3^X \Delta(x)^2\,dx, \tag{3}$$

where

$$\Delta(x) = \psi(x) - x = \sum_{n \le x} \Lambda(n) - x.$$

Here $\Lambda(n)$ is von Mangoldt's function, defined by $\Lambda(n) = \log p$ if $n = p^k$, and $\Lambda(n) = 0$ otherwise.
Results for (2) can be deduced from (3) by partial summation.

The predicted mean square error (assuming RH)

We can assume RH, since otherwise Cramér's model CM is false. CM predicts that

$$\frac{1}{X}\int_3^X \Delta(x)^2\,dx \sim X\log X. \tag{4}$$

On the other hand, Riemann's "explicit formula"

$$\Delta(x) = \lim_{T\to\infty}\sum_{|\rho|\le T}\frac{x^\rho}{\rho} + O(\log x),$$

where ρ runs over the nontrivial zeros of $\zeta(s)$, implies

$$\frac{1}{X}\int_3^X \Delta(x)^2\,dx = O(X). \tag{5}$$

This contradicts (4)!

Sketch of proof

Use a truncated version of the explicit formula

$$\Delta(x) = \sum_{|\rho| \le T} \frac{x^\rho}{\rho} + O(\log x) \quad \text{for} \quad T \ge x \log x.$$

Multiply out $\Delta(x)\overline{\Delta(x)}$ and integrate over $[3, X]$. This gives a sum of terms

$$\frac{X^{\rho+\rho'+1}}{\rho\rho'(\rho+\rho'+1)},$$

plus some lower-order terms which can be ignored. Here $\rho = \frac{1}{2} + i\gamma$ and $\rho' = \frac{1}{2} + i\gamma'$ are nontrivial zeros of $\zeta(s)$.
Split the sum into two, say S_1 over (ρ, ρ') such that

$$|\Im(\rho+\rho')| \le 1, \tag{6}$$

and S_2 over the remaining pairs (ρ, ρ'). Using the fact that, for each zero ρ, there are $O(\log|\Im(\rho)|)$ zeros ρ' such that (6) holds, we can prove that each sum is $O(X^2)$.

The key idea

Considering just the nontrivial zeros ρ_n in the upper half-plane, we have $\rho_n = 1/2 + i\gamma_n$, where $\gamma_n \sim 2\pi n/\log n$; more precisely

$$n = \frac{\gamma_n}{2\pi}\left(\log\left(\frac{\gamma_n}{2\pi}\right) - 1\right) + O(\log \gamma_n).$$

Thus,

$$\frac{|S_1|}{X^2} \ll \sum_{n=1}^{O(T)} \frac{(\log n)^3}{n^2},$$

and

$$\frac{|S_2|}{X^2} \ll \sum_{n=1}^{O(T)} \frac{\log n}{n} \sum_{n'=1}^{n} \frac{\log n'}{n'(n-n'+1)} \ll \sum_{n=1}^{O(T)} \frac{(\log n)^3}{n(n+1)}.$$

However, both these series are absolutely convergent.

Some constants

Assuming RH, we obtain

$$c_0 \le \limsup_{X\to\infty} \frac{1}{X^2} \int_3^X \Delta(x)^2\,dx \le c_1,$$

where the lower bound comes from the terms with $\rho' = \overline{\rho}$ (and assumes that the γ_n are independent over $\mathbb{Q}$):

$$c_0 = 4 \sum_{n=1}^{\infty} \frac{1}{1+4\gamma_n^2} \approx 0.023.$$

The upper bound comes from bounding all the terms (not best possible as the terms can not all have phase close to zero):

$$c_1 = 8 \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{[4+(\gamma_m-\gamma_n)^2]^{-1/2} + [4+(\gamma_m+\gamma_n)^2]^{-1/2}}{[(1+4\gamma_m^2)(1+4\gamma_n^2)]^{1/2}} \approx 0.21.$$

Typical values are much smaller than c_1. For $X \in [10^6, 10^7]$, the maximum is 0.0296 near $X = 3609803$.

Generalisation

Using the same ideas, we can prove

$$\frac{1}{X}\int_3^X |\Delta(x)|^\alpha \, dx = O_\alpha(X^{\alpha/2}) \tag{7}$$

for all positive integers α.

However, $\max_{3\le x\le X} |\Delta(x)| \neq O(X^{1/2})$.

Littlewood (1914) proved the lower bound

$$\Delta(X) = \Omega_\pm(X^{1/2}\log_3 X).$$

This does not contradict (7), because the worst case occurs only rarely.

A related problem

Soundararajan, in his lecture notes for a 2006 Summer School, discusses a closely related problem. Assuming RH, he outlines a proof that, when h/X is small,

$$\frac{1}{X}\int_X^{2X} (\psi(x+h) - \psi(x) - h)^2 \, \mathrm{d}x \ll h(1 + (\log X)/h)^2. \qquad (8)$$

A similar result was proved by Selberg (1943).

Assume that $h/\log X$ is large, but h/X is small. CM predicts that $\psi(X+h) - \psi(X)$ is approximately normal with mean $\sim h$ and variance $\sim h \log X$. The CM prediction for the variance disagrees with (8) by a factor of order $\log X$.

Putting $h = X$ formally gives our result (although Soundararajan's proof may not apply as h/X is not small).

Numerical results (A)

To confirm the theory for primes, and to see what happens for twin primes (where we do not have any theoretical results), we performed some statistical tests.

Define Chebyshev's function

$$\theta_1(x) := \sum_{\text{primes } p \le x} \log p$$

and analogously

$$\theta_2(x) := \sum_{\text{twins } (p,p+2),\ p \le x} \log(p)\log(p+2).$$

On Cramér's model we would expect $(\theta_1(x) - x)^2$ to be of order $x \log x$, more precisely to have mean

$$V_1(x) = (x + 1/2)(\log x - 2) + O(1);$$

and $(\theta_2(x) - 2c_2 x)^2$ to be of order $x \log^2 x$, more precisely

$$V_2(x) = 2c_2(x + 3/2)(\log x - 1)^2 + O(\log x).$$

x	$\frac{(\psi(x)-x)^2}{x}$	$\frac{(\theta_1(x)-x)^2}{x}$	$\frac{(\theta_1(x)-x)^2}{V_1(x)}$	$\frac{(\theta_2(x)-x)^2}{V_2(x)}$
10^2	0.355	2.65	0.203	1.56
10^3	0.011	1.91	0.055	1.85
10^4	0.018	1.08	0.016	0.07
10^5	0.027	0.99	0.009	0.22
10^6	0.171	2.30	0.014	0.57
10^7	0.213	2.32	0.010	1.08
10^8	0.031	1.50	0.005	0.02

Table: Test of Cramér's model for primes and twin primes

The numerical results confirm that both $\psi(x) - x$ and $\theta_1(x) - x$ have mean square values of order x, although $|\psi(x) - x|$ is usually[2] smaller than $|\theta_1(x) - x|$.

As expected, CM makes incorrect predictions for primes, but the numerical evidence for twins is inconclusive.

[2] Not always smaller, by the result of Littlewood (1914).

Numerical results (B)

Consider the interval $I = [n_0, n_0 + n)$. We split I into ν bins, each of size n/ν. For each bin B_k, we count the number of primes O_k in the bin. We compute the expected number of primes E_k using

$$E_k = \sum_{2 \le m \in B_k} \frac{1}{\log m} .$$

Similarly for twin-prime pairs $(m, m+2)$, except that now

$$E_k = \sum_{3 \le m \in B_k} \frac{2c_2}{\log(m)\log(m+2)} .$$

Numerical results

In each case we compute the "chi-squared" statistic

$$\chi_\nu^2 := \sum_{k=1}^{\nu} \frac{(O_k - E_k)^2}{E_k} \, .$$

On the null hypothesis (which corresponds to Cramér's model being true), we expect χ_ν^2 to have a chi-squared distribution with ν degrees of freedom (provided that the bins are sufficiently large so E_k is not too small).

We can compute the p-value (the probability that an observed χ_ν^2 would be exceeded), and a q-value ($q = 1 - p$).

If p is small, e.g. $p < 0.01$, then the prediction is poor and the null hypothesis can be rejected.

On the other hand, if q is small, then the prediction is too good! In both cases the null hypothesis should be rejected.

Numerical results

n_0	n	ν	q (primes)	q (twins)
0	10^6	100	2.1×10^{-16}	3.3×10^{-4}
0	10^7	1000	2.4×10^{-96}	4.5×10^{-18}
10^7	10^7	1000	1.2×10^{-76}	5.3×10^{-11}
10^8	10^7	1000	1.5×10^{-58}	5.3×10^{-8}
10^9	10^7	1000	2.4×10^{-35}	3.7×10^{-2}
10^{10}	10^7	1000	8.7×10^{-31}	1.3×10^{-4}

Table: Test of Cramér's model for primes and twin primes

We see that neither the primes nor the twin primes satisfy the null hypothesis; the prediction is too good.

The q-values for twins are larger than the corresponding q-values for primes; in this sense the twins are "more random" than the primes.

Some history

Pintz (2007) states the results (4)–(5) and mentions that he found them “a few years after Maier’s (1985) discovery” but did not publish them at the time.

Pintz mentions that a result equivalent to (5) was proved by Cramér in 1920, about 15 years before the publication of Cramér’s model.

Thus, we see the reason for the title of Pintz’s paper:

“Cramér vs Cramér”.

It is not clear whether Cramér was aware of the contradiction between his model and his earlier result.


Dustin Hoffman
Kramer
vs.
Kramer

References

Gunnar Blom, Harald Cramér 1893–1985, *The Annals of Statistics* 15, 4 (1987), 1335–1350.

Richard Brent, Irregularities in the distribution of primes and twin primes, *Math. Comp.* 29 (1975), 43–56.

Harald Cramér, On the distribution of primes, *Proc. Camb. Phil. Soc.* 20 (1920), 272–280.

Harald Cramér, Some theorems concerning prime numbers, *Arkiv f. Math. Astr. Fys.* 15, 5 (1920), 1–33. (Thanks to Adrian Dudek for helping me obtain a copy of this paper!)

Harald Cramér, Prime numbers and probability, *Proc. Eighth Scand Math. Congress*, Stockholm (1935), 107–115.

Harald Cramér, On the order of magnitude of the difference between consecutive prime numbers, *Acta Arith.* 2 (1936), 23–46.

P. X. Gallagher, On the distribution of primes in short intervals, *Mathematika* 23 (1976), 4–9.

Andrew Granville, Harald Cramér and the distribution of prime numbers, *Scand. Actuarial J.* 1995, #1, 12–28.

Andrew Granville and Kannan Soundararajan, An uncertainty principle for arithmetic sequences, *Annals of Math.* 165 (2007), 593–635.

G. H. Hardy and J. E. Littlewood, Some problems on partitio numerorum III: On the expression of a number a s a sum of primes, *Acta Math.* 44 (1923), 1–70.

Helmut Maier, Primes in short intervals, *Michigan Math. J.* 32 (1985), 221–225.

János Pintz, Cramér vs. Cramér: on Cramér's probabilistic model for primes, *Functiones et Approximatio* XXXVII.2 (2007), 361–376.

Atle Selberg, On the normal density of primes in short intervals, and the difference between consecutive primes, *Arch. Math. Naturvid.* 47 (1943), 87–105.

Kannan Soundararajan, *The distribution of prime numbers*, Lectures at NATO School on Equidistribution, arXiv:math/0606408v1, 16 June 2006.