The distribution of consecutive prime biases

Robert J. Lemke Oliver [1]
Tufts University

(Joint with K. Soundararajan)

[1] Partially supported by NSF grant DMS-1601398

Chebyshev's Bias

Let $\pi(x; q, a) := \#\{p < x : p \equiv a (\text{mod } q)\}$.

Chebyshev's Bias

Let $\pi(x; q, a) := \#\{p < x : p \equiv a (\text{mod } q)\}$. Under GRH, we have

$$\pi(x; q, a) = \frac{\mathrm{li}(x)}{\phi(q)} + O(x^{1/2+\epsilon}) \quad \text{if } (a, q) = 1.$$

Chebyshev's Bias

Let $\pi(x; q, a) := \#\{p < x : p \equiv a (\text{mod } q)\}$. Under GRH, we have

$$\pi(x; q, a) = \frac{\mathrm{li}(x)}{\phi(q)} + O(x^{1/2+\epsilon}) \quad \text{if } (a, q) = 1.$$

Observation (Chebyshev)

Quadratic residues seem slightly less frequent than non-residues.

Chebyshev's Bias

Let $\pi(x; q, a) := \#\{p < x : p \equiv a (\text{mod } q)\}$. Under GRH, we have

$$\pi(x; q, a) = \frac{\mathrm{li}(x)}{\phi(q)} + O(x^{1/2+\epsilon}) \quad \text{if } (a, q) = 1.$$

Observation (Chebyshev)

Quadratic residues seem slightly less frequent than non-residues.

Theorem (Rubinstein-Sarnak)

Under GRH$(+\epsilon)$, $\pi(x; 3, 2) > \pi(x; 3, 1)$ *for* 99.9% *of* x,

Chebyshev's Bias

Let $\pi(x; q, a) := \#\{p < x : p \equiv a (\text{mod } q)\}$. Under GRH, we have

$$\pi(x; q, a) = \frac{\mathrm{li}(x)}{\phi(q)} + O(x^{1/2+\epsilon}) \quad \text{if } (a, q) = 1.$$

Observation (Chebyshev)

Quadratic residues seem slightly less frequent than non-residues.

Theorem (Rubinstein-Sarnak)

Under GRH($+\epsilon$), $\pi(x; 3, 2) > \pi(x; 3, 1)$ for 99.9% of x, and analogous results hold for any q.

The primes (mod 3)

$\pi(x)$	$\pi(x;3,1)$	$\pi(x;3,2)$
10^3	491	508

The primes (mod 3)

$\pi(x)$	$\pi(x;3,1)$	$\pi(x;3,2)$
10^3	491	508
10^4	4989	5010

The primes (mod 3)

$\pi(x)$	$\pi(x;3,1)$	$\pi(x;3,2)$
10^3	491	508
10^4	4989	5010
10^5	49962	50037

The primes (mod 3)

$\pi(x)$	$\pi(x;3,1)$	$\pi(x;3,2)$
10^3	491	508
10^4	4989	5010
10^5	49962	50037
10^6	499829	500170

The primes (mod 3)

$\pi(x)$	$\pi(x;3,1)$	$\pi(x;3,2)$
10^3	491	508
10^4	4989	5010
10^5	49962	50037
10^6	499829	500170
10^7	4999505	5000494

The primes (mod 4)

$\pi(x)$	$\pi(x; 4, 1)$	$\pi(x; 4, 3)$

The primes (mod 4)

$\pi(x)$	$\pi(x;4,1)$	$\pi(x;4,3)$
10^3	495	504
10^4	4984	5015
10^5	49950	50049
10^6	499798	500201
10^7	4999452	5000547

The primes (mod 5)

$\pi(x)$	$\pi(x;5,1)$	$\pi(x;5,2)$	$\pi(x;5,3)$	$\pi(x;5,4)$

The primes (mod 5)

$\pi(x)$	$\pi(x;5,1)$	$\pi(x;5,2)$	$\pi(x;5,3)$	$\pi(x;5,4)$
10^3	245	256	253	246

The primes (mod 5)

$\pi(x)$	$\pi(x;5,1)$	$\pi(x;5,2)$	$\pi(x;5,3)$	$\pi(x;5,4)$
10^3	245	256	253	246
10^4	2485	2508	2515	2492

The primes (mod 5)

$\pi(x)$	$\pi(x;5,1)$	$\pi(x;5,2)$	$\pi(x;5,3)$	$\pi(x;5,4)$
10^3	245	256	253	246
10^4	2485	2508	2515	2492
10^5	24968	25015	25008	25009

The primes (mod 5)

$\pi(x)$	$\pi(x;5,1)$	$\pi(x;5,2)$	$\pi(x;5,3)$	$\pi(x;5,4)$
10^3	245	256	253	246
10^4	2485	2508	2515	2492
10^5	24968	25015	25008	25009
10^6	249934	250017	250109	249940

The primes (mod 5)

$\pi(x)$	$\pi(x;5,1)$	$\pi(x;5,2)$	$\pi(x;5,3)$	$\pi(x;5,4)$
10^3	245	256	253	246
10^4	2485	2508	2515	2492
10^5	24968	25015	25008	25009
10^6	249934	250017	250109	249940
10^7	2499756	2500284	2500208	2499752

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$,

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

We expect that $\pi(x; q, \mathbf{a}) \sim \mathrm{li}(x)/\phi(q)^r$,

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

We expect that $\pi(x; q, \mathbf{a}) \sim \text{li}(x)/\phi(q)^r$, but little is known:

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

We expect that $\pi(x; q, \mathbf{a}) \sim \mathrm{li}(x)/\phi(q)^r$, but little is known:

- If $r = 2$ and $\phi(q) = 2$, then each $\mathbf{a}$ occurs infinitely often

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

We expect that $\pi(x; q, \mathbf{a}) \sim \mathrm{li}(x)/\phi(q)^r$, but little is known:

- If $r = 2$ and $\phi(q) = 2$, then each $\mathbf{a}$ occurs infinitely often
- Shiu: The pattern $(a, a, \ldots, a)$ occurs infinitely often

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

We expect that $\pi(x; q, \mathbf{a}) \sim \mathrm{li}(x)/\phi(q)^r$, but little is known:

- If $r = 2$ and $\phi(q) = 2$, then each $\mathbf{a}$ occurs infinitely often
- Shiu: The pattern $(a, a, \ldots, a)$ occurs infinitely often
- Maynard: $\pi(x; q, (a, a, \ldots, a)) \gg \pi(x)$

Patterns of consecutive primes

Let $r \geq 1$ and $\mathbf{a} = (a_1, \ldots, a_r)$, and set

$$\pi(x; q, \mathbf{a}) := \#\{p_n < x : p_{n+i} \equiv a_{i+1} (\text{mod } q) \text{ for } 0 \leq i < r\}.$$

We expect that $\pi(x; q, \mathbf{a}) \sim \mathrm{li}(x)/\phi(q)^r$, but little is known:

- If $r = 2$ and $\phi(q) = 2$, then each $\mathbf{a}$ occurs infinitely often
- Shiu: The pattern $(a, a, \ldots, a)$ occurs infinitely often
- Maynard: $\pi(x; q, (a, a, \ldots, a)) \gg \pi(x)$

Question

Are there biases between the different patterns $\mathbf{a} (\text{mod } q)$?

Consecutive primes (mod 3)

$\pi(x)$	$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$	$\pi(x;3,(2,1))$	$\pi(x;3,(2,2))$

Consecutive primes (mod 3)

$\pi(x)$	$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$	$\pi(x;3,(2,1))$	$\pi(x;3,(2,2))$
10^3	179	313	313	195

Consecutive primes (mod 3)

$\pi(x)$	$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$	$\pi(x;3,(2,1))$	$\pi(x;3,(2,2))$
10^3	179	313	313	195
10^4	1999	2990	2991	2020

Consecutive primes (mod 3)

$\pi(x)$	$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$	$\pi(x;3,(2,1))$	$\pi(x;3,(2,2))$
10^3	179	313	313	195
10^4	1999	2990	2991	2020
10^5	20933	29029	29030	21008

Consecutive primes (mod 3)

$\pi(x)$	$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$	$\pi(x;3,(2,1))$	$\pi(x;3,(2,2))$
10^3	179	313	313	195
10^4	1999	2990	2991	2020
10^5	20933	29029	29030	21008
10^6	215873	283957	283957	216213

Consecutive primes (mod 3)

$\pi(x)$	$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$	$\pi(x;3,(2,1))$	$\pi(x;3,(2,2))$
10^3	179	313	313	195
10^4	1999	2990	2991	2020
10^5	20933	29029	29030	21008
10^6	215873	283957	283957	216213
10^7	2203296	2796210	2796210	2204284

Consecutive primes (mod 4)

$\pi(x)$	$\pi(x; 4, (1,1))$	$\pi(x; 4, (1,3))$	$\pi(x; 4, (3,1))$	$\pi(x; 4, (3,3))$
10^3	187	309	309	195
10^4	2053	2931	2931	2085
10^5	21269	28681	28680	21370
10^6	218510	281289	281288	218913
10^7	2223704	2775749	2775748	2224799

Consecutive primes (mod 5)

Let $\pi(x_0) = 10^7$.

Consecutive primes (mod 5)

Let $\pi(x_0) = 10^7$. We find:

a	$\pi(x_0; 5, a)$
1	2,499,755
2	2,500,283

a	$\pi(x_0; 5, a)$
3	2,500,209
4	2,499,751

Consecutive primes (mod 5)

Let $\pi(x_0) = 10^7$. We find:

a	b	$\pi(x_0; 5, (a, b))$
1	1	2,499,755
	2	
	3	
	4	
	2	2,500,283

a	$\pi(x_0; 5, a)$
3	2,500,209
4	2,499,751

Consecutive primes (mod 5)

Let $\pi(x_0) = 10^7$. We find:

a	b	$\pi(x_0; 5, (a, b))$
1	1	446,808
	2	769,923
	3	756,071
	4	526,953
	2	2,500,283

a	$\pi(x_0; 5, a)$
3	2,500,209
4	2,499,751

Consecutive primes (mod 5)

Let $\pi(x_0) = 10^7$. We find:

a	b	$\pi(x_0; 5, (a, b))$
1	1	446,808
	2	769,923
	3	756,071
	4	526,953
2	1	639,384
	2	422,289
	3	681,759
	4	756,851

a	b	$\pi(x_0; 5, (a, b))$
3	1	593,195
	2	714,795
	3	422,302
	4	769,915
4	1	820,368
	2	593,275
	3	640,076
	4	446,032

Consecutive primes (mod 5)

Let $\pi(x_1) = 10^8$. We find:

a	b	$\pi(x_1; 5, (a, b))$
1	1	4,623,041
	2	7,504,612
	3	7,429,438
	4	5,442,344
2	1	6,373,982
	2	4,439,355
	3	6,755,195
	4	7,431,870

a	b	$\pi(x_1; 5, (a, b))$
3	1	6,010,981
	2	7,043,695
	3	4,442,561
	4	7,502,896
4	1	7,991,431
	2	6,012,739
	3	6,372,940
	4	4,622,916

What's the deal?

What's the deal?

We conjecture that:

- There are large secondary terms in the asymptotic for $\pi(x; q, \mathbf{a})$

What's the deal?

We conjecture that:

- There are large secondary terms in the asymptotic for $\pi(x; q, \mathbf{a})$
- The dominant factor is the number of $a_i \equiv a_{i+1} (\text{mod } q)$

What's the deal?

We conjecture that:

- There are large secondary terms in the asymptotic for $\pi(x; q, \mathbf{a})$
- The dominant factor is the number of $a_i \equiv a_{i+1} \pmod{q}$
- There are smaller, somewhat erratic factors that affect non-diagonal $\mathbf{a}$

The conjecture: explicit version

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$*.*

The conjecture: explicit version

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$. *Then*

$$\pi(x; q, \mathbf{a}) = \frac{\mathrm{li}(x)}{\phi(q)^r} \left[1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + \frac{c_2(q; \mathbf{a})}{\log x} + O\left(\log^{-7/4} x \right) \right],$$

The conjecture: explicit version

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$. *Then*

$$\pi(x; q, \mathbf{a}) = \frac{\mathrm{li}(x)}{\phi(q)^r} \left[1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + \frac{c_2(q; \mathbf{a})}{\log x} + O\left(\log^{-7/4} x \right) \right],$$

where

$$c_1(q; \mathbf{a}) = \frac{\phi(q)}{2} \left(\frac{r-1}{\phi(q)} - \#\{1 \leq i < r : a_i \equiv a_{i+1} (mod\ q)\} \right),$$

The conjecture: explicit version

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$. *Then*

$$\pi(x; q, \mathbf{a}) = \frac{\mathrm{li}(x)}{\phi(q)^r} \left[1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + \frac{c_2(q; \mathbf{a})}{\log x} + O\left(\log^{-7/4} x \right) \right],$$

where

$$c_1(q; \mathbf{a}) = \frac{\phi(q)}{2} \left(\frac{r-1}{\phi(q)} - \#\{1 \leq i < r : a_i \equiv a_{i+1} (mod\ q)\} \right),$$

and $c_2(q; \mathbf{a})$ *is complicated but explicit.*

An example

Example

Let $q = 3$ or 4.

An example

Example

Let $q = 3$ or 4. Then

$$\pi(x; q, (a, b)) = \frac{\mathrm{li}(x)}{4}\left[1 \pm \left(\frac{\log\log x}{2\log x} + \frac{\log 2\pi/q}{2\log x}\right)\right] + O\left(\frac{x}{\log^{11/4} x}\right).$$

An example

Example

Let $q = 3$ or 4. Then

$$\pi(x;q,(a,b)) = \frac{\mathrm{li}(x)}{4}\left[1 \pm \left(\frac{\log\log x}{2\log x} + \frac{\log 2\pi/q}{2\log x}\right)\right] + O\left(\frac{x}{\log^{11/4} x}\right).$$

Conjecture (LO & Soundararajan)

Let $q = 3$ or 4. If $a \not\equiv b(\text{mod } q)$, then for all $x > 5$,

$$\pi(x;q,(a,b)) > \pi(x;q,(a,a)).$$

Comparison with numerics (mod 3)

Comparison with numerics (mod 3)

x		$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$
10^9	Actual	$1.132 \cdot 10^7$	$1.411 \cdot 10^7$

Comparison with numerics (mod 3)

x		$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$
10^9	Actual	$1.132 \cdot 10^7$	$1.411 \cdot 10^7$
	Conj.	$1.137 \cdot 10^7$	$1.405 \cdot 10^7$

Comparison with numerics (mod 3)

x		$\pi(x;3,(1,1))$	$\pi(x;3,(1,2))$
10^9	Actual	$1.132 \cdot 10^7$	$1.411 \cdot 10^7$
	Conj.	$1.137 \cdot 10^7$	$1.405 \cdot 10^7$
10^{10}		$1.024 \cdot 10^8$	$1.251 \cdot 10^8$
		$1.028 \cdot 10^8$	$1.247 \cdot 10^8$
10^{11}		$9.347 \cdot 10^8$	$1.124 \cdot 10^9$
		$9.383 \cdot 10^8$	$1.121 \cdot 10^9$
10^{12}		$8.600 \cdot 10^9$	$1.020 \cdot 10^{10}$
		$8.630 \cdot 10^9$	$1.017 \cdot 10^{10}$

The conjecture when $q = 5$

When $q = 5$,

The conjecture when $q = 5$

When $q = 5$, we predict that

$$c_2(5; (a, a)) = \frac{-3 \log(2\pi/5)}{2},$$

The conjecture when $q = 5$

When $q = 5$, we predict that

$$c_2(5; (a, a)) = \frac{-3 \log(2\pi/5)}{2},$$

and if $a \neq b$, then $c_2(5; (a, b)) =$

$$\frac{\log(2\pi/5)}{2} + \frac{5}{2}\Re\left(L(0, \chi)L(1, \chi)A_{5,\chi}\left[\bar{\chi}(b - a) + \frac{\bar{\chi}(b) - \bar{\chi}(a)}{4}\right]\right),$$

The conjecture when $q = 5$

When $q = 5$, we predict that

$$c_2(5; (a,a)) = \frac{-3\log(2\pi/5)}{2},$$

and if $a \neq b$, then $c_2(5; (a,b)) =$

$$\frac{\log(2\pi/5)}{2} + \frac{5}{2}\Re\left(L(0,\chi)L(1,\chi)A_{5,\chi}\left[\bar{\chi}(b-a) + \frac{\bar{\chi}(b) - \bar{\chi}(a)}{4}\right]\right),$$

where

$$A_{5,\chi} = \prod_{p \neq 5}\left(1 - \frac{(\chi(p)-1)^2}{(p-1)^2}\right) \approx 1.891 + 1.559i.$$

Comparison with numerics: $q = 5$

x	$\pi(x; 5, (1,1))$	$\pi(x; 5, (1,2))$	$\pi(x; 5, (1,3))$	$\pi(x; 5, (1,4))$
10^9	$2.328 \cdot 10^6$ $2.354 \cdot 10^6$	$3.842 \cdot 10^6$ $3.774 \cdot 10^6$	$3.796 \cdot 10^6$ $3.835 \cdot 10^6$	$2.745 \cdot 10^6$ $2.750 \cdot 10^6$

Comparison with numerics: $q = 5$

x	$\pi(x;5,(1,1))$	$\pi(x;5,(1,2))$	$\pi(x;5,(1,3))$	$\pi(x;5,(1,4))$
10^9	$2.328 \cdot 10^6$ $2.354 \cdot 10^6$	$3.842 \cdot 10^6$ $3.774 \cdot 10^6$	$3.796 \cdot 10^6$ $3.835 \cdot 10^6$	$2.745 \cdot 10^6$ $2.750 \cdot 10^6$
10^{10}	$2.142 \cdot 10^7$ $2.164 \cdot 10^7$	$3.369 \cdot 10^7$ $3.324 \cdot 10^7$	$3.348 \cdot 10^7$ $3.374 \cdot 10^7$	$2.516 \cdot 10^7$ $2.515 \cdot 10^7$

Comparison with numerics: $q = 5$

x	$\pi(x;5,(1,1))$	$\pi(x;5,(1,2))$	$\pi(x;5,(1,3))$	$\pi(x;5,(1,4))$
10^9	$2.328 \cdot 10^6$	$3.842 \cdot 10^6$	$3.796 \cdot 10^6$	$2.745 \cdot 10^6$
	$2.354 \cdot 10^6$	$3.774 \cdot 10^6$	$3.835 \cdot 10^6$	$2.750 \cdot 10^6$
10^{10}	$2.142 \cdot 10^7$	$3.369 \cdot 10^7$	$3.348 \cdot 10^7$	$2.516 \cdot 10^7$
	$2.164 \cdot 10^7$	$3.324 \cdot 10^7$	$3.374 \cdot 10^7$	$2.515 \cdot 10^7$
10^{11}	$1.984 \cdot 10^8$	$3.000 \cdot 10^8$	$2.993 \cdot 10^8$	$2.318 \cdot 10^8$
	$2.002 \cdot 10^8$	$2.969 \cdot 10^8$	$3.011 \cdot 10^8$	$2.314 \cdot 10^8$

Comparison with numerics: $q = 5$

x	$\pi(x;5,(1,1))$	$\pi(x;5,(1,2))$	$\pi(x;5,(1,3))$	$\pi(x;5,(1,4))$
10^9	$2.328 \cdot 10^6$	$3.842 \cdot 10^6$	$3.796 \cdot 10^6$	$2.745 \cdot 10^6$
	$2.354 \cdot 10^6$	$3.774 \cdot 10^6$	$3.835 \cdot 10^6$	$2.750 \cdot 10^6$
10^{10}	$2.142 \cdot 10^7$	$3.369 \cdot 10^7$	$3.348 \cdot 10^7$	$2.516 \cdot 10^7$
	$2.164 \cdot 10^7$	$3.324 \cdot 10^7$	$3.374 \cdot 10^7$	$2.515 \cdot 10^7$
10^{11}	$1.984 \cdot 10^8$	$3.000 \cdot 10^8$	$2.993 \cdot 10^8$	$2.318 \cdot 10^8$
	$2.002 \cdot 10^8$	$2.969 \cdot 10^8$	$3.011 \cdot 10^8$	$2.314 \cdot 10^8$
10^{12}	$1.848 \cdot 10^9$	$2.704 \cdot 10^9$	$2.706 \cdot 10^9$	$2.145 \cdot 10^9$
	$1.863 \cdot 10^9$	$2.682 \cdot 10^9$	$2.717 \cdot 10^9$	$2.141 \cdot 10^9$

The heuristic when $r = 2$

$$\pi(x; q, (a, b)) =$$

The heuristic when $r = 2$

$$\pi(x; q, (a, b)) = \sum_{\substack{n<x: \\ n\equiv a \pmod{q}}} \mathbf{1}_{\mathcal{P}}(n)$$

The heuristic when $r = 2$

$$\pi(x; q, (a, b)) = \sum_{\substack{n<x: \\ n \equiv a \pmod{q}}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0: \\ h \equiv b-a \pmod{q}}} \mathbf{1}_{\mathcal{P}}(n+h)$$

The heuristic when $r = 2$

$$\pi(x; q, (a, b)) = \sum_{\substack{n<x: \\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0: \\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) \cdot$$

$$\cdot \prod_{\substack{t<h: \\ (t+a,q)=1}} \Big(1 - \mathbf{1}_{\mathcal{P}}(n+t)\Big)$$

The heuristic when $r = 2$

$$\pi(x;q,(a,b)) \approx \sum_{\substack{n<x:\\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0:\\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) \cdot$$

$$\cdot \prod_{\substack{t<h:\\ (t+a,q)=1}} \left(1 - \frac{1}{\log x}\right)$$

Please do not try this at home!

The heuristic when $r = 2$

$$\pi(x; q, (a, b)) \approx \sum_{\substack{n<x: \\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0: \\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) \cdot$$

$$\cdot \prod_{\substack{t<h: \\ (t+a,q)=1}} \left(1 - \frac{q}{\phi(q)} \frac{1}{\log x}\right)$$

The heuristic when $r = 2$

$$\pi(x;q,(a,b)) \approx \sum_{\substack{n<x:\\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0:\\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) \cdot$$

$$\cdot \prod_{\substack{t<h:\\ (t+a,q)=1}} \Big(1 - \frac{q}{\phi(q)}\frac{1}{\log x}\Big)$$

$$\sim \sum_{\substack{n<x:\\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0:\\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) e^{-h/\log x}$$

The heuristic when $r = 2$

$$\pi(x; q, (a,b)) \approx \sum_{\substack{n<x:\\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0:\\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) \cdot$$

$$\cdot \prod_{\substack{t<h:\\ (t+a,q)=1}} \Big(1 - \frac{q}{\phi(q)} \frac{1}{\log x}\Big)$$

$$\sim \sum_{\substack{n<x:\\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \sum_{\substack{h>0:\\ h\equiv b-a \pmod q}} \mathbf{1}_{\mathcal{P}}(n+h) e^{-h/\log x}$$

$$= \sum_{h\equiv b-a \pmod q} e^{-h/\log x} \sum_{\substack{n<x:\\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n) \mathbf{1}_{\mathcal{P}}(n+h)$$

The Hardy-Littlewood conjecture

We need to understand

$$\sum_{h \equiv b-a \pmod q} e^{-h/\log x} \sum_{\substack{n<x \\ n \equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h).$$

The Hardy-Littlewood conjecture

We need to understand

$$\sum_{h \equiv b-a \pmod q} e^{-h/\log x} \sum_{\substack{n<x \\ n \equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h).$$

Conjecture (Hardy-Littlewood)

For any $h \neq 0$, we have

$$\sum_{n<x} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h) \sim \mathfrak{S}(h)\frac{x}{\log^2 x},$$

The Hardy-Littlewood conjecture

We need to understand

$$\sum_{h \equiv b-a (\mathrm{mod}\ q)} e^{-h/\log x} \sum_{\substack{n<x \\ n \equiv a (\mathrm{mod}\ q)}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h).$$

Conjecture (Hardy-Littlewood)

For any $h \neq 0$, we have

$$\sum_{n<x} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h) \sim \mathfrak{S}(h)\frac{x}{\log^2 x},$$

where

$$\mathfrak{S}(h) = \prod_p \left(1 - \frac{\#(\{0,h\} mod\, p)}{p}\right)\left(1-\frac{1}{p}\right)^{-2}.$$

The Hardy-Littlewood conjecture

We need to understand

$$\sum_{h \equiv b-a (\mathrm{mod}\ q)} e^{-h/\log x} \sum_{\substack{n<x \\ n \equiv a (\mathrm{mod}\ q)}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h).$$

Conjecture (Hardy-Littlewood)

For any $h \neq 0$, we have

$$\sum_{\substack{n<x \\ n \equiv a (mod\ q)}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h) \sim \mathfrak{S}(h)\frac{x}{\log^2 x},$$

where

$$\mathfrak{S}(h) = \prod_p \left(1 - \frac{\#(\{0,h\} mod\, p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-2}.$$

The Hardy-Littlewood conjecture

We need to understand

$$\sum_{h\equiv b-a \pmod q} e^{-h/\log x} \sum_{\substack{n<x \\ n\equiv a \pmod q}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h).$$

Conjecture (Hardy-Littlewood)

For any $h \neq 0$ with $(h+a,q)=1$, we have

$$\sum_{\substack{n<x \\ n\equiv a (mod\ q)}} \mathbf{1}_{\mathcal{P}}(n)\mathbf{1}_{\mathcal{P}}(n+h) \sim \frac{q}{\phi(q)^2}\mathfrak{S}_q(h)\frac{x}{\log^2 x},$$

where

$$\mathfrak{S}_q(h) = \prod_{p\nmid q}\left(1-\frac{\#(\{0,h\} mod\, p)}{p}\right)\left(1-\frac{1}{p}\right)^{-2}.$$

The main term

We now have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b - a \pmod{q}} \mathfrak{S}_q(h) e^{-h/\log x}.$$

The main term

We now have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b-a \pmod{q}} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Gallagher: $\mathfrak{S}_q(h) = 1$ on average

The main term

We now have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b - a (\text{mod } q)} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Gallagher: $\mathfrak{S}_q(h) = 1$ on average

Blindly plugging this in, we have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b - a (\text{mod } q)} e^{-h/\log x}$$

The main term

We now have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b - a (\text{mod } q)} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Gallagher: $\mathfrak{S}_q(h) = 1$ on average

Blindly plugging this in, we have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b - a (\text{mod } q)} e^{-h/\log x}$$
$$\sim \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \frac{\log x}{q}$$

The main term

We now have

$$\pi(x; q, (a,b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b-a (\mathrm{mod}\ q)} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Gallagher: $\mathfrak{S}_q(h) = 1$ on average

Blindly plugging this in, we have

$$\begin{aligned}\pi(x; q, (a,b)) &\approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b-a (\mathrm{mod}\ q)} e^{-h/\log x} \\ &\sim \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \frac{\log x}{q} = \frac{1}{\phi(q)^2} \frac{x}{\log x},\end{aligned}$$

The main term

We now have

$$\pi(x; q, (a, b)) \approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b-a (\mathrm{mod}\ q)} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Gallagher: $\mathfrak{S}_q(h) = 1$ on average

Blindly plugging this in, we have

$$\begin{aligned} \pi(x; q, (a, b)) &\approx \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \sum_{h \equiv b-a (\mathrm{mod}\ q)} e^{-h/\log x} \\ &\sim \frac{q}{\phi(q)^2} \frac{x}{\log^2 x} \frac{\log x}{q} = \frac{1}{\phi(q)^2} \frac{x}{\log x}, \end{aligned}$$

which is the main term.

The source of the bias

Consider

$$\sum_{h \equiv b-a \pmod{q}} \mathfrak{S}_q(h) e^{-h/\log x}.$$

The source of the bias

Consider

$$\sum_{h \equiv b-a \pmod{q}} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Proposition

We have

$$\sum_{h \equiv 0 (mod\ q)} \mathfrak{S}_q(h) e^{-h/\log x} = \frac{\log x}{q} - \frac{\phi(q)}{2q} \log\log x + O(1),$$

The source of the bias

Consider

$$\sum_{h \equiv b-a (\text{mod } q)} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Proposition

We have

$$\sum_{h \equiv 0 (mod\ q)} \mathfrak{S}_q(h) e^{-h/\log x} = \frac{\log x}{q} - \frac{\phi(q)}{2q} \log\log x + O(1),$$

while for $v \not\equiv 0 (mod\ q)$,

$$\sum_{h \equiv v (mod\ q)} \mathfrak{S}_q(h) e^{-h/\log x} = \frac{\log x}{q} + O(1).$$

The source of the bias

Consider

$$\sum_{h \equiv b-a (\text{mod } q)} \mathfrak{S}_q(h) e^{-h/\log x}.$$

Proposition

We have

$$\sum_{h \equiv 0 (mod\ q)} \mathfrak{S}_q(h) e^{-h/\log x} = \frac{\log x}{q} - \frac{\phi(q)}{2q} \log\log x + O(1),$$

while for $v \not\equiv 0 (mod\ q)$*,*

$$\sum_{h \equiv v (mod\ q)} \mathfrak{S}_q(h) e^{-h/\log x} = \frac{\log x}{q} + O(1).$$

Idea

Only the first Dirichlet series has a pole at $s = 0$.

Much needed rigor

To do this properly, we need to be more careful with

$$\prod_{\substack{t<h:\\(t+a,q)=1}} (1-\mathbf{1}_{\mathcal{P}}(n+t)) \approx \prod_{\substack{t<h:\\(t+a,q)=1}} \left(1-\frac{q}{\phi(q)\log x}\right).$$

Much needed rigor

To do this properly, we need to be more careful with

$$\prod_{\substack{t<h:\\(t+a,q)=1}} (1-\mathbf{1}_{\mathcal{P}}(n+t)) \approx \prod_{\substack{t<h:\\(t+a,q)=1}} \left(1-\frac{q}{\phi(q)\log x}\right).$$

Could expand out and invoke Hardy-Littlewood,

Much needed rigor

To do this properly, we need to be more careful with

$$\prod_{\substack{t<h:\\(t+a,q)=1}} (1 - \mathbf{1}_{\mathcal{P}}(n+t)) \approx \prod_{\substack{t<h:\\(t+a,q)=1}} \left(1 - \frac{q}{\phi(q)\log x}\right).$$

Could expand out and invoke Hardy-Littlewood, but this is ugly!

Much needed rigor

To do this properly, we need to be more careful with

$$\prod_{\substack{t<h: \\ (t+a,q)=1}} (1 - \mathbf{1}_{\mathcal{P}}(n+t)) \approx \prod_{\substack{t<h: \\ (t+a,q)=1}} \left(1 - \frac{q}{\phi(q)\log x}\right).$$

Could expand out and invoke Hardy-Littlewood, but this is ugly!

Better idea: Incorporate inclusion-exclusion directly into H-L.

Modified Hardy-Littlewood

Define $\widetilde{\mathbf{1}}_{\mathcal{P}}(n) = \mathbf{1}_{\mathcal{P}}(n) - \frac{q}{\phi(q)\log n}$.

Modified Hardy-Littlewood

Define $\widetilde{\mathbf{1}}_{\mathcal{P}}(n) = \mathbf{1}_{\mathcal{P}}(n) - \frac{q}{\phi(q)\log n}$.

Conjecture

If $|\mathcal{H}| = k$ *with* $(h+a, q) = 1$ *for all* $h \in \mathcal{H}$*, then*

$$\sum_{\substack{n \leq x \\ n \equiv a (mod\ q)}} \prod_{h \in \mathcal{H}} \widetilde{\mathbf{1}}_{\mathcal{P}}(n+h) \sim \frac{q^{k-1}}{\phi(q)^k} \mathfrak{S}_{q,0}(\mathcal{H}) \frac{x}{\log^k x},$$

Modified Hardy-Littlewood

Define $\widetilde{\mathbf{1}}_{\mathcal{P}}(n) = \mathbf{1}_{\mathcal{P}}(n) - \frac{q}{\phi(q)\log n}$.

Conjecture

If $|\mathcal{H}| = k$ *with* $(h+a,q)=1$ *for all* $h \in \mathcal{H}$, *then*

$$\sum_{\substack{n \le x \\ n \equiv a (mod\ q)}} \prod_{h \in \mathcal{H}} \widetilde{\mathbf{1}}_{\mathcal{P}}(n+h) \sim \frac{q^{k-1}}{\phi(q)^k} \mathfrak{S}_{q,0}(\mathcal{H}) \frac{x}{\log^k x},$$

where

$$\mathfrak{S}_{q,0}(\mathcal{H}) := \sum_{\mathcal{T} \subseteq \mathcal{H}} (-1)^{|\mathcal{H} \setminus \mathcal{T}|} \mathfrak{S}_q(\mathcal{T}).$$

Sums of modified singular series

Theorem (Montgomery, Soundararajan)

$$\sum_{\substack{\mathcal{H}\subseteq[1,h]\\|\mathcal{H}|=k}} \mathfrak{S}_0(\mathcal{H}) = \frac{\mu_k}{k!}(-h\log h + Ah)^{k/2} + O_k(h^{k/2-\delta}),$$

where $\mu_k = 0$ *if* k *is odd.*

Sums of modified singular series

Theorem (Montgomery, Soundararajan)

$$\sum_{\substack{\mathcal{H}\subseteq[1,h]\\ |\mathcal{H}|=k}} \mathfrak{S}_0(\mathcal{H}) = \frac{\mu_k}{k!}(-h\log h + Ah)^{k/2} + O_k(h^{k/2-\delta}),$$

where $\mu_k = 0$ *if* k *is odd.*

Point

We can discard $\mathcal{H}$ with $|\mathcal{H}| \geq 3$.

Notes on assembly

Three flavors of two-term $\mathcal{H}$:

Notes on assembly

Three flavors of two-term $\mathcal{H}$:

- $\mathcal{H} = \{0, h\}$: Contributes main-term, bias against diagonal

Notes on assembly

Three flavors of two-term $\mathcal{H}$:

- $\mathcal{H} = \{0, h\}$: Contributes main-term, bias against diagonal
- $\mathcal{H} = \{0, t\}$, $\{t, h\}$: (One interloper) Contributes to c_2

Notes on assembly

Three flavors of two-term $\mathcal{H}$:

- $\mathcal{H} = \{0, h\}$: Contributes main-term, bias against diagonal
- $\mathcal{H} = \{0, t\}$, $\{t, h\}$: (One interloper) Contributes to c_2
- $\mathcal{H} = \{t_1, t_2\}$: (Two interlopers) Unbiased

Notes on assembly

Three flavors of two-term $\mathcal{H}$:

- $\mathcal{H} = \{0, h\}$: Contributes main-term, bias against diagonal
- $\mathcal{H} = \{0, t\}$, $\{t, h\}$: (One interloper) Contributes to c_2
- $\mathcal{H} = \{t_1, t_2\}$: (Two interlopers) Unbiased

Remark

We expect $\mathcal{H}$ with $|\mathcal{H}| \geq 3$ to contribute further lower-order terms.

The conjecture

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$. *Then*

$$\pi(x; q, \mathbf{a}) = \frac{\mathrm{li}(x)}{\phi(q)^r} \left[1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + \frac{c_2(q; \mathbf{a})}{\log x} + O\left(\log^{-7/4} x \right) \right],$$

where

$$c_1(q; \mathbf{a}) = \frac{\phi(q)}{2} \left(\frac{r-1}{\phi(q)} - \#\{1 \leq i < r : a_i \equiv a_{i+1} (mod\ q)\} \right),$$

and $c_2(q; \mathbf{a})$ *is complicated but explicit.*

The constant $c_2(q; \mathbf{a})$

The constant $c_2(q; \mathbf{a})$

If q is prime, then

$$c_2(q; (a, a)) = \frac{q-2}{2} \log(q/2\pi),$$

The constant $c_2(q; \mathbf{a})$

If q is prime, then

$$c_2(q; (a, a)) = \frac{q-2}{2} \log(q/2\pi),$$

while

$$\frac{c_2(q; (a, b))}{q} = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} \bar{\chi}(b-a) L(0, \chi) L(1, \chi) A_{q,\chi} + o(1),$$

The constant $c_2(q; \mathbf{a})$

If q is prime, then

$$c_2(q;(a,a)) = \frac{q-2}{2}\log(q/2\pi),$$

while

$$\frac{c_2(q;(a,b))}{q} = \frac{1}{\phi(q)} \sum_{\chi\neq\chi_0 (\mathrm{mod}\ q)} \bar{\chi}(b-a)L(0,\chi)L(1,\chi)A_{q,\chi} + o(1),$$

where

$$A_{q,\chi} = \prod_{p\nmid q}\left(1 - \frac{(\chi(p)-1)^2}{(p-1)^2}\right).$$

How big is $c_2(q; \mathbf{a})$? A heuristic

How big is $c_2(q; \mathbf{a})$? A heuristic

Expect $L(1, \chi) \approx 1$ and $L(0, \chi) \approx q^{1/2}$ for "typical" χ,

How big is $c_2(q; \mathbf{a})$? A heuristic

Expect $L(1,\chi) \approx 1$ and $L(0,\chi) \approx q^{1/2}$ for "typical" χ, so roughly

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi}$$

How big is $c_2(q;\mathbf{a})$? A heuristic

Expect $L(1,\chi)\approx 1$ and $L(0,\chi)\approx q^{1/2}$ for "typical" χ, so roughly

$$\frac{1}{\phi(q)}\sum_{\chi\neq\chi_0 \pmod q}\bar{\chi}(b-a)L(0,\chi)L(1,\chi)A_{q,\chi}$$

$$\approx \frac{q^{1/2}}{\phi(q)}\sum_{\chi\neq\chi_0 \pmod q} e(\theta_\chi)$$

How big is $c_2(q;\mathbf{a})$? A heuristic

Expect $L(1,\chi) \approx 1$ and $L(0,\chi) \approx q^{1/2}$ for "typical" χ, so roughly

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi}$$

$$\approx \frac{q^{1/2}}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} e(\theta_\chi)$$

$$\approx \frac{q^{1/2}}{\phi(q)} \cdot q^{1/2}$$

How big is $c_2(q; \mathbf{a})$? A heuristic

Expect $L(1,\chi) \approx 1$ and $L(0,\chi) \approx q^{1/2}$ for "typical" χ, so roughly

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod q} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi}$$

$$\begin{aligned} &\approx \frac{q^{1/2}}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod q} e(\theta_\chi) \\ &\approx \frac{q^{1/2}}{\phi(q)} \cdot q^{1/2} \\ &\approx 1. \end{aligned}$$

How big is $c_2(q;\mathbf{a})$? A heuristic

Expect $L(1,\chi) \approx 1$ and $L(0,\chi) \approx q^{1/2}$ for "typical" χ, so roughly

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi}$$

$$\approx \frac{q^{1/2}}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} e(\theta_\chi)$$

$$\approx \frac{q^{1/2}}{\phi(q)} \cdot q^{1/2}$$

$$\approx 1.$$

Thus, we should typically expect $c_2(q;\mathbf{a})/q \approx 1$.

The distribution of $c_2(q; \mathbf{a})$

Theorem (LO & Soundararajan)

As $q \to \infty$, $c_2(q;(a,b))/q$ *has a continuous limiting distribution,*

The distribution of $c_2(q; \mathbf{a})$

Theorem (LO & Soundararajan)

As $q \to \infty$, $c_2(q; (a, b))/q$ has a continuous limiting distribution, e.g.

$$\lim_{q\to\infty} \frac{\#\{a \not\equiv b (mod\ q) : c_2(q; (a, b))/q \leq \tau\}}{\phi(q)(\phi(q) - 1)}$$

exists and is non-zero.

The distribution of $c_2(q; \mathbf{a})$

Theorem (LO & Soundararajan)

As $q \to \infty$, $c_2(q; (a, b))/q$ has a continuous limiting distribution, e.g.

$$\lim_{q\to\infty} \frac{\#\{a \not\equiv b (mod\ q) : c_2(q; (a, b))/q \le \tau\}}{\phi(q)(\phi(q) - 1)}$$

exists and is non-zero. Moreover, there is $A > 0$ such that

$$\mathbb{P}\left[\left|\frac{c_2(q; (a, b))}{q}\right| > \tau\right] \le \exp(-Ae^{\tau}/\tau).$$

The distribution of $c_2(q; \mathbf{a})$

Theorem (LO & Soundararajan)

As $q \to \infty$, $c_2(q; (a, b))/q$ has a continuous limiting distribution, e.g.

$$\lim_{q\to\infty} \frac{\#\{a \not\equiv b (mod\ q) : c_2(q; (a, b))/q \leq \tau\}}{\phi(q)(\phi(q) - 1)}$$

exists and is non-zero. Moreover, there is $A > 0$ such that

$$\mathbb{P}\left[\left|\frac{c_2(q; (a, b))}{q}\right| > \tau\right] \leq \exp(-Ae^{\tau}/\tau).$$

Remark

This is reminiscent of the distribution of $L(1, \chi)$.

A simplified version

Recall that

$$\frac{c_2(q;(a,b))}{q} = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(b-a)L(0,\chi)L(1,\chi)A_{q,\chi} + o(1).$$

A simplified version

Recall that

$$\frac{c_2(q;(a,b))}{q} = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi} + o(1).$$

Consider instead the simplified version

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t) L(0,\chi) L(1,\chi).$$

A simplified version

Recall that

$$\frac{c_2(q;(a,b))}{q} = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi} + o(1).$$

Consider instead the simplified version

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} \bar{\chi}(t) L(0,\chi) L(1,\chi).$$

It turns out this is related to the Dedekind sum $s_q(a)$,

A simplified version

Recall that

$$\frac{c_2(q;(a,b))}{q} = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} \bar{\chi}(b-a) L(0,\chi) L(1,\chi) A_{q,\chi} + o(1).$$

Consider instead the simplified version

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\text{mod } q)} \bar{\chi}(t) L(0,\chi) L(1,\chi).$$

It turns out this is related to the Dedekind sum $s_q(a)$, given by

$$s_q(a) = \sum_{x (\text{mod } q)}{}' \psi(x/q)\psi(ax/q), \quad \psi(x) = x - \lfloor x \rfloor - 1/2.$$

Dedekind sums

The Dedekind sum $s_q(a)$ is

- large when a/q is close to a rational with small denominator,

Dedekind sums

The Dedekind sum $s_q(a)$ is

- large when a/q is close to a rational with small denominator, e.g.

$$s_q(1) = \frac{q}{12} - \frac{1}{4} + \frac{1}{6q},$$

Dedekind sums

The Dedekind sum $s_q(a)$ is

- large when a/q is close to a rational with small denominator, e.g.
$$s_q(1) = \frac{q}{12} - \frac{1}{4} + \frac{1}{6q}, \text{ but}$$
- it's typically much smaller:

Dedekind sums

The Dedekind sum $s_q(a)$ is

- large when a/q is close to a rational with small denominator, e.g.
$$s_q(1) = \frac{q}{12} - \frac{1}{4} + \frac{1}{6q}, \text{ but}$$
- it's typically much smaller:

Theorem (Vardi)

As $q \to \infty$, $\frac{s_q(a)}{\log q}$ is dist. according to the Cauchy distribution.

Dedekind sums

The Dedekind sum $s_q(a)$ is

- large when a/q is close to a rational with small denominator, e.g.
$$s_q(1) = \frac{q}{12} - \frac{1}{4} + \frac{1}{6q}, \text{ but}$$
- it's typically much smaller:

Theorem (Vardi)

As $q \to \infty$, $\frac{s_q(a)}{\log q}$ is dist. according to the Cauchy distribution.

Remark

$d\mu_{Cauchy} = \frac{2}{1+4\pi^2x^2}\,dx$ has infinite expectation.

The Fourier transform of Dedekind sums

Consider for $0 \leq t \leq q-1$ the Fourier transform

$$\hat{s}_q(t) := \frac{1}{q} \sum_{a \pmod{q}} s_q(a) e(at/q).$$

The Fourier transform of Dedekind sums

Consider for $0 \leq t \leq q-1$ the Fourier transform

$$\hat{s}_q(t) := \frac{1}{q} \sum_{a \pmod q} s_q(a) e(at/q).$$

Then

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod q} \bar{\chi}(t) L(0,\chi) L(1,\chi)$$

The Fourier transform of Dedekind sums

Consider for $0 \leq t \leq q-1$ the Fourier transform

$$\hat{s}_q(t) := \frac{1}{q} \sum_{a \pmod{q}} s_q(a) e(at/q).$$

Then

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t) L(0,\chi) L(1,\chi) = -\pi i \cdot \hat{s}_q(t).$$

The Fourier transform of Dedekind sums

Consider for $0 \le t \le q-1$ the Fourier transform

$$\hat{s}_q(t) := \frac{1}{q} \sum_{a \pmod{q}} s_q(a) e(at/q).$$

Then

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t) L(0,\chi) L(1,\chi) = -\pi i \cdot \hat{s}_q(t).$$

Theorem (LO & Soundararajan)

As $q \to \infty$, $\hat{s}_q(t)$ has a continuous limiting distribution.

A closer connection to $L(1, \chi)$

Using

$$L(1, \chi) = \sum_{b \geq 1} \frac{\chi(b)}{b}$$

and

$$L(0, \chi) = \frac{-1}{q} \sum_{a \pmod{q}} \chi(a)\psi(a/q)$$

with orthogonality of characters,

A closer connection to $L(1,\chi)$

Using

$$L(1,\chi) = \sum_{b\geq 1} \frac{\chi(b)}{b}$$

and

$$L(0,\chi) = \frac{-1}{q} \sum_{a \pmod q} \chi(a)\psi(a/q)$$

with orthogonality of characters, we find

$$\frac{1}{\phi(q)} \sum_{\chi\neq\chi_0 \pmod q} \bar{\chi}(t) L(0,\chi) L(1,\chi)$$

A closer connection to $L(1, \chi)$

Using

$$L(1,\chi) = \sum_{b \geq 1} \frac{\chi(b)}{b}$$

and

$$L(0,\chi) = \frac{-1}{q} \sum_{a (\mathrm{mod}\ q)} \chi(a)\psi(a/q)$$

with orthogonality of characters, we find

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\mathrm{mod}\ q)} \bar{\chi}(t) L(0,\chi) L(1,\chi) = - \sum_{\substack{b \geq 1 \\ (b,q)=1}} \frac{\psi(t\bar{b}/q)}{b},$$

A closer connection to $L(1,\chi)$

Using

$$L(1,\chi) = \sum_{b \geq 1} \frac{\chi(b)}{b}$$

and

$$L(0,\chi) = \frac{-1}{q} \sum_{a (\mathrm{mod}\ q)} \chi(a)\psi(a/q)$$

with orthogonality of characters, we find

$$\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 (\mathrm{mod}\ q)} \bar{\chi}(t) L(0,\chi) L(1,\chi) = - \sum_{\substack{b \geq 1 \\ (b,q)=1}} \frac{\psi(t\bar{b}/q)}{b},$$

where $\bar{b}$ is the inverse of $b (\mathrm{mod}\ q)$.

The random model

Thus, we've found

$$\hat{s}_q(t) = \frac{1}{\pi i} \sum_{\substack{b \geq 1 \\ (b,q)=1}} \frac{\psi(t\bar{b}/q)}{b}.$$

The random model

Thus, we've found

$$\hat{s}_q(t) = \frac{1}{\pi i} \sum_{\substack{b \geq 1 \\ (b,q)=1}} \frac{\psi(t\bar{b}/q)}{b}.$$

Idea: Compare this to the random model

$$\frac{1}{\pi i} \sum_{b \geq 1} \frac{\psi(x/b)}{b}, \quad x \in \mathbb{R},$$

The random model

Thus, we've found

$$\hat{s}_q(t) = \frac{1}{\pi i} \sum_{\substack{b \geq 1 \\ (b,q)=1}} \frac{\psi(t\bar{b}/q)}{b}.$$

Idea: Compare this to the random model

$$\frac{1}{\pi i} \sum_{b \geq 1} \frac{\psi(x/b)}{b}, \quad x \in \mathbb{R},$$

and use the method of moments.

The random model

Thus, we've found

$$\hat{s}_q(t) = \frac{1}{\pi i} \sum_{\substack{b \geq 1 \\ (b,q)=1}} \frac{\psi(t\bar{b}/q)}{b}.$$

Idea: Compare this to the random model

$$\frac{1}{\pi i} \sum_{b \geq 1} \frac{\psi(x/b)}{b}, \quad x \in \mathbb{R},$$

and use the method of moments.

Remark

A similar expression holds for $c_2(q; \mathbf{a})$.

Another connection: A conjecture of Montgomery

Another connection: A conjecture of Montgomery

Conjecture (Montgomery)

Define $R(x)$ by

$$\sum_{n \le x} \phi(n) = \frac{3}{\pi^2} x^2 + R(x).$$

Another connection: A conjecture of Montgomery

Conjecture (Montgomery)

Define $R(x)$ by

$$\sum_{n \le x} \phi(n) = \frac{3}{\pi^2} x^2 + R(x).$$

Then $R(x) \ll x \log\log x$ and $R(x) = \Omega_{\pm}(x \log\log x)$.

Another connection: A conjecture of Montgomery

Conjecture (Montgomery)

Define $R(x)$ *by*

$$\sum_{n \le x} \phi(n) = \frac{3}{\pi^2} x^2 + R(x).$$

Then $R(x) \ll x \log\log x$ *and* $R(x) = \Omega_{\pm}(x \log\log x)$*.*

Theorem (LO & Soundararajan)

$R(x)/x$ *has a limiting distribution with doubly exponential decay.*

Wait, how is this related?

Wait, how is this related?

Theorem (Montgomery)

$R(x) = \Omega_{\pm}(x\sqrt{\log \log x})$.

Wait, how is this related?

Theorem (Montgomery)
$R(x) = \Omega_{\pm}(x\sqrt{\log\log x})$.

Idea: Montgomery shows that

$$R(x) = x \sum_{b \le x} \frac{\mu(b)\psi(x/b)}{b} + O(x/\log x),$$

Wait, how is this related?

Theorem (Montgomery)

$R(x) = \Omega_{\pm}(x\sqrt{\log\log x})$.

Idea: Montgomery shows that

$$R(x) = x \sum_{b \leq x} \frac{\mu(b)\psi(x/b)}{b} + O(x/\log x),$$

and chooses x very cleverly.

Wait, how is this related?

Theorem (Montgomery)
$R(x) = \Omega_{\pm}(x\sqrt{\log\log x})$.

Idea: Montgomery shows that

$$R(x) = x \sum_{b \leq x} \frac{\mu(b)\psi(x/b)}{b} + O(x/\log x),$$

and chooses x very cleverly.

But this is very close to our random model!

The conjecture

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$. *Then*

$$\pi(x; q, \mathbf{a}) = \frac{\mathrm{li}(x)}{\phi(q)^r} \left[1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + \frac{c_2(q; \mathbf{a})}{\log x} + O\left(\log^{-7/4} x\right)\right],$$

where

$$c_1(q; \mathbf{a}) = \frac{\phi(q)}{2} \left(\frac{r-1}{\phi(q)} - \#\{1 \leq i < r : a_i \equiv a_{i+1} (mod\ q)\}\right),$$

and $c_2(q; \mathbf{a})$ *is complicated, explicit,*

The conjecture

Conjecture (LO & Soundararajan)

Let $\mathbf{a} = (a_1, \ldots, a_r)$ *with* $r \geq 2$. *Then*

$$\pi(x; q, \mathbf{a}) = \frac{\mathrm{li}(x)}{\phi(q)^r} \left[1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + \frac{c_2(q; \mathbf{a})}{\log x} + O\left(\log^{-7/4} x \right) \right],$$

where

$$c_1(q; \mathbf{a}) = \frac{\phi(q)}{2} \left(\frac{r-1}{\phi(q)} - \#\{1 \leq i < r : a_i \equiv a_{i+1} (mod\ q)\} \right),$$

and $c_2(q; \mathbf{a})$ *is complicated, explicit, and nicely distributed.*