How unfair are prime number races?

Greg Martin
University of British Columbia

Québec–Vermont Number Theory Seminar
Centre de recherches mathématiques
Université de Montréal
February 2, 2006

For any positive real number V, let X_V denote a random variable having a normal distribution with mean 0 and variance V, so that

$$\Pr(X_V < x) = \frac{1}{\sqrt{2\pi V}} \int_{-\infty}^{x} e^{-t^2/2V}\, dt.$$

Then if the three random variables X_A, X_B, X_C are independent,

$$\Pr(X_A > X_B > X_C) =$$

For any positive real number V, let X_V denote a random variable having a normal distribution with mean 0 and variance V, so that

$$\Pr(X_V < x) = \frac{1}{\sqrt{2\pi V}} \int_{-\infty}^{x} e^{-t^2/2V}\, dt.$$

Then if the three random variables X_A, X_B, X_C are independent,

$$\Pr(X_A > X_B > X_C) = \frac{1}{2\pi} \tan^{-1} \frac{\sqrt{AB + BC + CA}}{B}\,.$$

For any positive real number V, let X_V denote a random variable having a normal distribution with mean 0 and variance V, so that

$$\Pr(X_V < x) = \frac{1}{\sqrt{2\pi V}} \int_{-\infty}^{x} e^{-t^2/2V}\, dt.$$

Then if the three random variables X_A, X_B, X_C are independent,

$$\Pr(X_A > X_B > X_C) = \frac{1}{2\pi} \tan^{-1} \frac{\sqrt{AB + BC + CA}}{B}.$$

One can check that

$$\Pr(X_A > X_B > X_C) + \Pr(X_A > X_C > X_B) + \Pr(X_B > X_A > X_C) \\ + \Pr(X_B > X_C > X_A) + \Pr(X_C > X_A > X_B) + \Pr(X_C > X_B > X_A) = 1$$

using the high school trigonometry formula

$$\tan^{-1}\alpha + \tan^{-1}\beta + \tan^{-1}\gamma \equiv \tan^{-1}\left(\frac{\alpha + \beta + \gamma - \alpha\beta\gamma}{1 - (\alpha\beta + \beta\gamma + \gamma\alpha)}\right) \pmod{\pi}.$$

How unfair are prime number races?

- What is “Chebyshev’s bias”?
- 20th-century technology
- My current research on two-way races
  - — Limiting distributions and their variance
  - — “Unfairness measure”: dependence on modulus
  - — “Unfairness measure”: dependence on residue classes
- My current(ish) research on three-way races
- Final sound bite (two-way races)

For relatively prime integers a and q, we define

$$\pi(x; q, a) = \#\{p \leq x \colon p \text{ prime}, \ p \equiv a \pmod q\}.$$

Chebyshev wrote a letter to M. Fuss in 1853 in which he said:

> *There is a notable difference in the splitting of the prime numbers between the two forms* $4n+3$, $4n+1$*: The first form contains a lot more than the second.*

Let's take various values of x and compare $\pi(x; 4, 3)$ to $\pi(x; 4, 1)$.

3	7	11	19	23

5	13	17	29

$$\pi(30;4,3) = 5 > 4 = \pi(30;4,1)$$

3	7	11	19	23	31	43	47	59

5	13	17	29	37	41	53

$$\pi(60;4,3) = 9 > 7 = \pi(60;4,1)$$

3	7	11	19	23	31	43	47	59	67	71	79	83

5	13	17	29	37	41	53	61	73	89

$$\pi(\ 90;4,3) = 13 > 10 = \pi(\ 90;4,1)$$

3	7	11	19	23	31	43	47	59	67	71	79	83	103	107

5	13	17	29	37	41	53	61	73	89	97	101	109

$$\pi(120;4,3) = 15 > 13 = \pi(120;4,1)$$

3	7	11	19	23	31	43	47	59	67	71	79	83	103	107	127	131	139

5	13	17	29	37	41	53	61	73	89	97	101	109	113	137

$$\pi(150;4,3) = 18 > 15 = \pi(150;4,1)$$

Races where such advantages are observed:

- Primes that are 2 (mod 3) over primes that are 1 (mod 3)
- Primes that are 3 (mod 4) over primes that are 1 (mod 4)
- Primes that are 2 or 3 (mod 5) over primes that are 1 or 4 (mod 5)
- Primes that are 3, 5, or 6 (mod 7) over primes that are 1, 2, or 4 (mod 7)
- Primes that are 3, 5, or 7 (mod 8) over primes that are 1 (mod 8)
- ...

The general pattern:

Primes that are nonsquares (mod q) over primes that are squares (mod q)

Why should the nonsquares have all the luck?

Let's look at an analogous situation, namely the comparison of $\pi(x)$ to $\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}$. The analytic proofs of the prime number theorem naturally give

$$\pi(x) + \tfrac{1}{2}\pi(x^{1/2}) + \tfrac{1}{3}\pi(x^{1/3}) + \cdots = \mathrm{li}(x) + \mathit{error},$$

where the error is roughly of size $\sqrt{x}$ if the Riemann Hypothesis is true. Therefore only the first two terms on the left are significant, and we derive

$$\pi(x) = \mathrm{li}(x) - \tfrac{1}{2}\,\mathrm{li}(x^{1/2}) + \mathit{error}.$$

If we divide through by $\sqrt{x}/\log x$ and put in the explicit form of the error, we obtain

$$\frac{\pi(x) - \mathrm{li}(x)}{\sqrt{x}/\log x} = -1 - \sum_{\gamma \in \mathbb{R}:\ \zeta(1/2+i\gamma)=0} \frac{e^{i\gamma \log x}}{1/2 + i\gamma} + o(1).$$

In the same way, the analytic proof of the prime number theorem for arithmetic progressions naturally gives

$$\pi(x;q,a) + \tfrac{1}{2} \sum_{\substack{b \pmod q \\ b^2 \equiv a \pmod q}} \pi(x^{1/2},q,b) + \cdots = \frac{\mathrm{li}(x)}{\phi(q)} + error,$$

which converts into

$$\frac{\phi(q)\pi(x;q,a) - \mathrm{li}(x)}{\sqrt{x}/\log x} = -\#\{b \pmod q \colon b^2 \equiv a \pmod q\}$$
$$- \frac{1}{\phi(q)} \sum_{\chi \pmod q} \overline{\chi(a)} \sum_{\gamma \in \mathbb{R}\colon L(1/2+i\gamma,\chi)=0} \frac{e^{i\gamma \log x}}{1/2+i\gamma} + o(1).$$

Therefore the residue classes that are squares modulo q have the deck stacked against them.

How unfair are prime number races?

- What is "Chebyshev's bias"?
- 20th-century technology
- My current research on two-way races
  - — Limiting distributions and their variance
  - — "Unfairness measure": dependence on modulus
  - — "Unfairness measure": dependence on residue classes
- My current(ish) research on three-way races
- Final sound bite (two-way races)

Main question we wanted answered last century:

> *Do these observed inequalities hold for all x? Or for all sufficiently large x?*
>
> *Or is this just an accident of small numbers, and the inequalities and their opposites are both equally likely in the long run?*

Further computation (1950s and beyond) reveals that there are moments of triumph for the square residue classes over nonsquare residue classes:

- $\pi(x;4,1) > \pi(x;4,3)$ for the first time at $x = 26{,}861$
  - However, 26,863 is also a prime, so $\pi(x;4,3)$ immediately draws even and doesn't give up the lead again until $x = 616{,}841$
- $\pi(x;8,1) > \pi(x;8,5)$ for the first time at $x = 588{,}067{,}889$ (and $\pi(x;8,1)$ has still never caught $\pi(x;8,3)$ or $\pi(x;8,7)$ up to that point)
- $\pi(x;3,1) > \pi(x;3,2)$ for the first time at $x = 608{,}981{,}813{,}029$

And theoretical results as well:

- The prime number theorem for arithmetic progressions ($1900 + O(1)$): $\pi(x; q, a) \sim \mathrm{li}(x)/\phi(q) \sim \pi(x; q, b)$
- Littlewood (1910s): each of $\pi(x; 4, 1)$ and $\pi(x; 4, 3)$ is ahead of the other for arbitrarily large x, and similarly for $\pi(x; 3, 1)$ and $\pi(x; 3, 2)$
- Turán and Knapowski (1960s): $\pi(x; q, a)$ is ahead of $\pi(x; q, b)$ for arbitrarily large x, for many pairs of residue classes a, b. However, assumptions on the locations of zeros of Dirichlet L-functions are appearing.
- Kaczorowski (1990s): further results in this vein and also for "3-way races", "4-way races", etc.

Even more ambitious question:

> *How can we attach a sensible number to these prime number races?*

In other words:

> **How often is $\pi(x; q, a)$ ahead of $\pi(x; q, b)$?**

Define $\delta_{q;a,b}$ to be the logarithmic density of the set of real numbers $x \ge 1$ satisfying $\pi(x;q,a) > \pi(x;q,b)$. More explicity,

$$\delta_{q;a,b} = \lim_{X\to\infty} \left(\frac{1}{\log X} \int\limits_{\substack{1\le x\le X \\ \pi(x;q,a)>\pi(x;q,b)}} \frac{dx}{x} \right).$$

Most important to remember:

> $\delta_{q;a,b}$ measures the limiting "probability" that when a "random" real number x is chosen, there are more primes that are congruent to $a \pmod q$ up to x then there are congruent to $b \pmod q$.

Rubinstein and Sarnak (1994) investigated these densities under the following two hypotheses:

- The Generalized Riemann Hypothesis (GRH): all nontrivial zeros of Dirichlet L-functions have real part equal to $\frac{1}{2}$

Note: Recent work of Ford and Konyagin shows that certain hypothetical violations of GRH do actually lead to pathological behavior in prime number races.

Rubinstein and Sarnak (1994) investigated these densities under the following two hypotheses:

- The Generalized Riemann Hypothesis (GRH): all nontrivial zeros of Dirichlet L-functions have real part equal to $\frac{1}{2}$
- A linear independence hypothesis (LI): the nonnegative imaginary parts of these nontrivial zeros are linearly independent over the rationals

Note: The linear independence hypothesis is somewhat analogous to a "nonsingularity" hypothesis: if we had precise information about any linear dependences that might exist, we could probably still work out the answer....

Under these two hypotheses, the Generalized Riemann Hypothesis (GRH) and the linear independence hypothesis (LI), Rubinstein and Sarnak proved:

- $\delta_{q;a,b}$ always exists and is strictly between 0 and 1
- $\delta_{q;a,b} + \delta_{q;b,a} = 1$
- $\delta_{q;a,b} > \frac{1}{2}$ if and only if a is a nonsquare $\pmod q$ and b is a square $\pmod q$
- if a and b are distinct squares $\pmod q$ or distinct nonsquares $\pmod q$, then $\delta_{q;a,b} = \delta_{q;b,a} = \frac{1}{2}$
- $\delta_{q;a,b}$ tends to $\frac{1}{2}$ as q tends to infinity, uniformly for all pairs a, b of distinct reduced residues $\pmod q$.

They also made some computations—for example, $\delta_{4;3,1} = 0.9959\dots$ and $\delta_{3;2,1} = 0.9990\dots$. In the $\pi(x)$ versus $\mathrm{li}(x)$ race, the corresponding bias towards $\mathrm{li}(x)$ is $0.99999973\dots$!

In joint work with A. Feuerverger (2000), we extended Rubinstein and Sarnak's approach in several directions. For example, we calculated (assuming, as usual, GRH and LI) many examples of the densities $\delta_{q;a,b}$.

One significant discovery is that even with q fixed, the values of $\delta_{q;a,b}$ vary significantly as a and b vary over squares and nonsquares (mod q).

We established some equalities between certain $\delta_{q;a,b}$:

- $\delta_{q;a,b} = \delta_{q;ab^{-1},1}$ for any square $b \pmod q$. Thus it suffices to calculate only the values of $\delta_{q;a,1}$ for nonsquares $a \pmod q$.
- $\delta_{q;a,1} = \delta_{q;a^{-1},1}$ for any $a \pmod q$, so even some of these densities are duplicated.

However, calculations show that otherwise the densities have distinct values:

Examples for the moduli $q = 24$ and $q = 43$:

a	$\delta_{24;a,1}$
5	0.999987
11	0.999983
23	0.999889
7	0.999833
19	0.999719
17	0.999125
13	0.998722

a	$a^{-1} \pmod{43}$	$\delta_{43;a,1} = \delta_{43;a^{-1},1}$
30	33	0.57044
32	39	0.57040
12	18	0.56904
20	28	0.56881
19	34	0.56613
8	27	0.56606
5	26	0.56366
7	37	0.56345
2	22	0.56281
3	29	0.56065
42	42	0.55982

Current goals:

• A more precise understanding of the sizes of $\delta_{q;a,b}$. Recalling that $\delta_{q;a,b}$ tends to $\frac{1}{2}$ as q tends to infinity, we would like an asymptotic formula for $\delta_{q;a,b} - \frac{1}{2}$, for example.

• A way to decide which $\delta_{q;a,b}$ are likely to be larger than others as a and b vary (with q fixed), based on elementary criteria rather than laborious numerical calculation.

• Better understanding of races among more than two residue classes

Everything hereafter will assume both GRH and LI.

How unfair are prime number races?

- What is “Chebyshev’s bias”?
- 20th-century technology
- My current research on two-way races

  — Limiting distributions and their variance

  — “Unfairness measure”: dependence on modulus

  — “Unfairness measure”: dependence on residue classes
- My current(ish) research on three-way races
- Final sound bite (two-way races)

Let a be a nonsquare $\pmod q$ and b be a square $\pmod q$. Under GRH,

$$\frac{\phi(q)(\pi(x;q,a)-\pi(x;q,b))}{\sqrt{x}/\log x} = \rho(q) + \sum_{\substack{\chi \pmod q \\ \chi \ne \chi_0}} \big(\overline{\chi(b)} - \overline{\chi(a)}\big) \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{e^{i\gamma \log x}}{1/2+i\gamma} + o(1),$$

where

$$\begin{aligned}\rho(q) &= \text{the number of square roots of 1} \pmod q \\ &= \text{the number of square roots of any square} \pmod q \\ &= \text{the number of quadratic characters} \pmod q \\ &= 2^{\#\text{number of odd prime factors of } q} \times \{1,\ 2,\ \text{or } 4\}.\end{aligned}$$

Let a be a nonsquare $\pmod q$ and b be a square $\pmod q$. Under GRH,

$$\frac{\phi(q)(\pi(x;q,a)-\pi(x;q,b))}{\sqrt{x}/\log x} = \rho(q) + \sum_{\chi \pmod q} \big(\overline{\chi(b)}-\overline{\chi(a)}\big) \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{e^{i\gamma\log x}}{1/2+i\gamma} + o(1).$$

Under LI, this has the same limiting (logarithmic) distribution as the random variable

$$\rho(q) + \sum_{\chi \pmod q} \big|\chi(b)-\chi(a)\big| \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{1/4+\gamma^2}},$$

where each X_γ has the "arcsin distribution" on $[-1,1]$ (that is, X_γ is the real part of a random variable uniformly distributed on the unit circle); and the various X_γ are all independent except that $X_{-\gamma}=X_\gamma$.

The mean of the random variable

$$\rho(q) + \sum_{\chi \pmod q} \bigl|\chi(b) - \chi(a)\bigr| \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{1/4+\gamma^2}}\,,$$

is obviously $\rho(q)$. Its variance, which we denote $V(q;a,b)$, will be quite important:

$$V(q;a,b) = \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \bigl|\chi(b) - \chi(a)\bigr|^2 b(\chi),$$

where

$$b(\chi) = \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{1/4+\gamma^2}.$$

The variance of the random variable

$$\rho(q) + \sum_{\chi \pmod q} \big|\chi(b) - \chi(a)\big| \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{1/4+\gamma^2}}\,,$$

which is also the variance of the limiting distribution of $\phi(q)(\pi(x;q,a) - \pi(x;q,b)) \log x/\sqrt{x}$, is given by

$$V(q;a,b) = \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \big|\chi(b) - \chi(a)\big|^2 b(\chi).$$

The constant $b(\chi)$ was understood classically: if χ is primitive to the modulus q, then $b(\chi) = \log q + O(\log\log q)$. It follows that

$$\begin{aligned} V(q;a,b) &\approx \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \big|\chi(b) - \chi(a)\big|^2 (\log q + O(\log\log q)) \\ &= 2\phi(q)\log q + O(\phi(q)\log\log q). \end{aligned}$$

How unfair are prime number races?

- What is “Chebyshev’s bias”?
- 20th-century technology
- My current research on two-way races
  - — Limiting distributions and their variance
  - — “Unfairness measure”: dependence on modulus
  - — “Unfairness measure”: dependence on residue classes
- My current(ish) research on three-way races
- Final sound bite (two-way races)

Let $f_{q;a,b}$ denote the limiting logarithmic distribution of $\phi(q)(\pi(x;q,a) - \pi(x;q,b)) \log x/\sqrt{x}$, which equals the distribution function of the random variable

$$\rho(q) + \sum_{\chi \pmod q} \bigl|\chi(b) - \chi(a)\bigr| \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{1/4+\gamma^2}}\,.$$

Rubinstein and Sarnak gave a formula for its Fourier transform:

$$\hat{f}_{q;a,b}(z) = e^{-i\rho(q)z} \prod_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \prod_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} J_0\biggl(\frac{2|\chi(b)-\chi(a)|z}{\sqrt{1/4+\gamma^2}}\biggr),$$

where J_0 is the Bessel function

$$J_0(t) = \sum_{m=0}^{\infty} \frac{(-1)^m (t/2)^{2m}}{(m!)^2} = 1 - \frac{t^2}{4} + O(t^4).$$

If we divide these random variables by their variances and take the limit:

(Central Limit) **Theorem**. Assume GRH and LI. Then the limiting logarithmic distribution functions of

$$\sqrt{\frac{\phi(q)}{2\log q}} \cdot \frac{\pi(x;q,a)-\pi(x;q,b)}{\sqrt{x}/\log x}$$

converge, as q tends to infinity (uniformly in a and b), to the standard normal distribution with mean 0 and variance 1.

By a 1925 theorem of Levy, it's enough to prove pointwise convergence of their Fourier transforms:

$$\lim_{q\to\infty} \hat{f}_{q;a,b}\left(\frac{\eta}{\sqrt{2\phi(q)\log q}}\right) = e^{-\eta^2/2}.$$

The proof begins by taking logarithms of both sides and using the Taylor expansion of $J_0(t)$ near $t=0$.

Let's remember why we care about the limiting distribution of $\phi(q)(\pi(x;q,a) - \pi(x;q,b)) \log x/\sqrt{x}$! The density $\delta_{q;a,b}$ is precisely the proportion of this distribution that corresponds to positive values.

The previous central limit theorem says that this limiting distribution resembles a normal distribution with mean $\rho(q)$ (assuming we've chosen a nonsquare a and a square b) and variance $V(q;a,b)$.

If it were *exactly* such a normal distribution, then the measure of the positive values would be exactly

$$\frac{1}{\sqrt{2\pi V(q;a,b)}} \int_{-\rho(q)}^{\infty} e^{-t^2/2V(q;a,b)}\, dt = \frac{1}{2} + \frac{1}{2}\mathrm{Erf}\left(\frac{\rho(q)}{\sqrt{2V(q;a,b)}}\right),$$

where

$$\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt = \frac{2x}{\sqrt{\pi}} + O(x^3).$$

Using quantitative formulations of the central limit theorem, we can in fact show that

$$\delta_{q;a,b} = \frac{1}{2} + \frac{1}{2}\,\mathrm{Erf}\left(\frac{\rho(q)}{\sqrt{2V(q;a,b)}}\right) + O\left(\frac{1}{\phi(q)\log q}\right).$$

And using the Taylor expansion $\mathrm{Erf}(x) = 2x/\sqrt{\pi} + O(x^3)$, we obtain:

Theorem. Assume GRH and LI. If a is a nonsquare $\pmod q$ and b is a square $\pmod q$, then

$$\delta_{q;a,b} = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} + O\left(\frac{1}{\phi(q)\log q}\right).$$

If a is a square $\pmod q$ and b is a nonsquare $\pmod q$, then

$$\delta_{q;a,b} = \frac{1}{2} - \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} + O\left(\frac{1}{\phi(q)\log q}\right).$$

We observed earlier that $V(q;a,b) \sim 2\phi(q)\log q$, and so we obtain a more straightforward corollary, but one with a weaker error term:

Corollary. Assume GRH and LI. If a is a nonsquare $\pmod q$ and b is a square $\pmod q$, then

$$\delta_{q;a,b} = \frac{1}{2} + \frac{\rho(q)}{2\sqrt{\pi}\,(\phi(q)\log q)^{1/2}} + O\Big(\frac{\rho(q)\log\log q}{\phi(q)^{1/2}(\log q)^{3/2}}\Big).$$

In particular, we have

$$\delta_{q;a,b} = \frac{1}{2} + O_\varepsilon\left(\frac{1}{q^{1/2-\varepsilon}}\right)$$

for any $\varepsilon > 0$.

How unfair are prime number races?

- What is “Chebyshev’s bias”?
- 20th-century technology
- My current research on two-way races

  — Limiting distributions and their variance

  — “Unfairness measure”: dependence on modulus

  — “Unfairness measure”: dependence on residue classes

- My current(ish) research on three-way races
- Final sound bite (two-way races)

Let's take a closer look at $V(q;a,b) = \sum_{\chi \pmod q} |\chi(b) - \chi(a)|^2 b(\chi)$.

The classical formula for $b(\chi)$, when χ is a primitive character (mod q), is

$$b(\chi) = \log \frac{q}{\pi} - \gamma_0 - (1 + \chi(-1)) \log 2 + 2\Re \frac{L'(1,\chi)}{L(1,\chi)},$$

where $\gamma_0 = \lim_{x\to\infty} \left(\sum_{n \le x} \frac{1}{n} - \log x \right) \approx 0.577216$ is Euler's constant.

Let's take a closer look at $V(q;a,b) = \sum_{\chi \pmod q} |\chi(b) - \chi(a)|^2 b(\chi)$.

The classical formula for $b(\chi)$, when χ is a primitive character $\pmod q$, is

$$b(\chi) = \log \frac{q}{\pi} - \gamma_0 - (1 + \chi(-1)) \log 2 + 2\Re \frac{L'(1,\chi)}{L(1,\chi)}.$$

Therefore

$$V(q;a,b) =$$

$$\sum_{\substack{d \mid q \\ d>1}} \sum_{\substack{\chi \pmod d \\ \chi \text{ primitive}}} \bigl|\chi(b) - \chi(a)\bigr|^2 \left(\log \frac{d}{\pi} - \gamma_0 - (1 + \chi(-1)) \log 2 + 2\Re \frac{L'(1,\chi)}{L(1,\chi)} \right).$$

The sum $V(q;a,b) =$

$$\sum_{\substack{d|q \\ d>1}} \sum_{\substack{\chi \pmod d \\ \chi \text{ primitive}}} \left|\chi(b) - \chi(a)\right|^2 \left(\log \frac{d}{\pi} - \gamma_0 - (1+\chi(-1)) \log 2 + 2\Re \frac{L'(1,\chi)}{L(1,\chi)} \right)$$

can be broken up into pieces, most of which are easy to evaluate using orthogonality of Dirichlet characters.

Example lemmas would be:

- $$\sum_{\substack{d|q \\ d>1}} \sum_{\substack{\chi \pmod d \\ \chi \text{ primitive}}} \log d = \phi(q) \left(\log q - \sum_{p|q} \frac{\log p}{p-1} \right)$$

- $$\sum_{\substack{d|q \\ d>1}} \sum_{\substack{\chi \pmod d \\ \chi \text{ primitive}}} \chi(a) \log d = -\phi(q) \frac{\Lambda(q/(q,a-1))}{\phi(q/(q,a-1))} \quad (a \not\equiv 1 \pmod q).$$

The result of these evaluations is:

Theorem. Assume GRH and LI. If a is a nonsquare $\pmod q$ and b is a square $\pmod q$, then

$$\delta_{q;a,b} = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} + O\Big(\frac{1}{\phi(q)\log q}\Big),$$

where

$$\begin{aligned} V(q;a,b) &= 2 \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} |\chi(b)-\chi(a)|^2 \sum_{\substack{\gamma>0 \\ L(\frac{1}{2}+i\gamma,\chi)=0}} \frac{1}{\frac{1}{4}+\gamma^2} \\ &= 2\phi(q)\bigg(\log q - \sum_{p\mid q}\frac{\log p}{p-1} - (\gamma_0 + \log 2\pi) + R_q(a-b)\bigg) \\ &\qquad + (2\log 2)\iota_q(-ab^{-1})\phi(q) + 2M(q;a,b). \end{aligned}$$

$$V(q;a,b) = 2\phi(q)\biggl(\log q - \sum_{p|q} \frac{\log p}{p-1} - (\gamma_0 + \log 2\pi) + R_q(a-b) \biggr) \\ + (2\log 2)\iota_q(-ab^{-1})\phi(q) + 2M(q;a,b).$$

There are three terms in this formula for the variance $V(q;a,b)$ that depend on a and b. Whenever any of the three is bigger than normal, the variance increases, causing the density $\delta_{q;a,b}$ to decrease.

Caveat 1: The variance $V(q;a,b)$ does not fully characterize the limiting distribution. So making predictions using only the variance is imperfect.

$$V(q;a,b) = 2\phi(q)\bigg(\log q - \sum_{p\mid q}\frac{\log p}{p-1} - (\gamma_0 + \log 2\pi) + R_q(a-b)\bigg) + (2\log 2)\iota_q(-ab^{-1})\phi(q) + 2M(q;a,b).$$

- $\iota_q(n) = \begin{cases} 1, & \text{if } n \equiv 1 \pmod{q}, \\ 0, & \text{if } n \not\equiv 1 \pmod{q} \end{cases}$
- $R_q(n) = \dfrac{\Lambda(q/(q,n))}{\phi(q/(q,n))}$
- $M(q;a,b) = \displaystyle\sum_{\substack{d\mid q \\ d>1}} \sum_{\substack{\chi \pmod{d} \\ \chi \text{ primitive}}} |\chi(a)-\chi(b)|^2 \frac{L'(1,\chi)}{L(1,\chi)}$

$$R_q(a-b) = \frac{\Lambda(q/(q,a-b))}{\phi(q/(q,a-b))}$$

provides extra variance (which reduces the corresponding density $\delta_{q;a,b}$) if a is congruent to b modulo appropriately large divisors of q:

a	$\delta_{24;a,1}$	$24/(24,a-1)$	$R_{24}(a-1)$
5		6	0
11		12	0
23		12	0
7		4	$(\log 2)/2$
19		4	$(\log 2)/2$
17		3	$(\log 3)/2$
13		2	$\log 2$

$$R_q(a-b) = \frac{\Lambda(q/(q,a-b))}{\phi(q/(q,a-b))}$$

provides extra variance (which reduces the corresponding density $\delta_{q;a,b}$) if a is congruent to b modulo appropriately large divisors of q:

a	$\delta_{24;a,1}$	$24/(24,a-1)$	$R_{24}(a-1)$
5	0.999987	6	0
11	0.999983	12	0
23	0.999889	12	0
7	0.999833	4	$(\log 2)/2$
19	0.999719	4	$(\log 2)/2$
17	0.999125	3	$(\log 3)/2$
13	0.998722	2	$\log 2$

$$\iota_q(-ab^{-1}) = \begin{cases} 1, & \text{if } -ab^{-1} \equiv 1 \pmod q, \\ 0, & \text{if } -ab^{-1} \not\equiv 1 \pmod q \end{cases}$$

$$M(q;a,b) = \sum_{\substack{d \mid q \\ d>1}} \sum_{\substack{\chi \pmod d \\ \chi \text{ primitive}}} |\chi(a) - \chi(b)|^2 \frac{L'(1,\chi^*)}{L(1,\chi^*)}$$

$\iota_q(-ab^{-1})$ provides extra variance exactly when $a \equiv -b \pmod q$. It can be shown that $M(q;a,b)$ tends to provide extra variance when there are small prime powers congruent to ab^{-1} and/or ba^{-1} modulo q.

Caveat 2: $M(q;a,b)$ is hard to pin down. It's the least precisely understood contribution to the variance.

$\iota_q(-ab^{-1})$ provides extra variance exactly when $a \equiv -b \pmod{q}$, while $M(q;a,b)$ tends to provide extra variance when there are small prime powers congruent to ab^{-1} and/or ba^{-1} modulo q.

a	$a^{-1} \pmod{43}$	$\delta_{43;a,1} = \delta_{43;a^{-1},1}$
30	33	
32	39	
12	18	
20	28	
19	34	
8	27	
5	26	
7	37	
2	22	
3	29	
42	42	

$\iota_q(-ab^{-1})$ provides extra variance exactly when $a \equiv -b \pmod q$, while $M(q;a,b)$ tends to provide extra variance when there are small prime powers congruent to ab^{-1} and/or ba^{-1} modulo q.

a	$a^{-1} \pmod{43}$	$\delta_{43;a,1} = \delta_{43;a^{-1},1}$
30	33	0.57044
32	39	0.57040
12	18	0.56904
20	28	0.56881
19	34	0.56613
8	27	0.56606
5	26	0.56366
7	37	0.56345
2	22	0.56281
3	29	0.56065
42	42	0.55982

Outreach: In a 1983 paper, Bays and Hudson graphed the prime counting functions modulo 11, normalized (essentially) in our familiar way:

$$\frac{\phi(q)\pi(x;11,a) - \mathrm{li}(x)}{\sqrt{x}/\log x}, \quad 1 \le a \le 10.$$

As expected, they found that first place rotated among $a = 2, 6, 7, 8, 10$, while last place rotated among $a = 1, 3, 4, 5, 9$.

However, quite surprisingly, they observed that:

- when $a = 2$ was in first place, $a = 9$ tended to be in last place
- when $a = 6$ was in first place, $a = 5$ tended to be in last place
- when $a = 7$ was in first place, $a = 4$ tended to be in last place
- when $a = 8$ was in first place, $a = 3$ tended to be in last place
- when $a = 10$ was in first place, $a = 1$ tended to be in last place

To say that $\big(\phi(q)\pi(x;11,a) - \mathrm{li}(x)\big)\log x/\sqrt{x}$ and $\big(\phi(q)\pi(x;11,11-a) - \mathrm{li}(x)\big)\log x/\sqrt{x}$ tend to be "mirror images" of each other is the same thing as saying that their sum is close to constant.

We can study more generally the functions

$$\frac{\phi(q)(\pi(q;x,a)+\pi(q;x,b)) - 2\,\mathrm{li}(x)}{\sqrt{x}/\log x}$$

where $q \equiv 3 \pmod 4$ is a prime, a is a nonsquare $\pmod q$, and b is a square $\pmod q$ and try to see why the choice $b = q - a$ should result in a more constant function than other choices.

The key observation is that "more constant" should correspond to a smaller variance.

All such functions have the same mean, and their variance can be calculated in a similar manner:

$$\begin{aligned} V_+(q;a,b) &= \sum_{\substack{\chi \pmod q \\ \chi\ne\chi_0}} |\chi(a)+\chi(b)|^2 b(\chi) \\ &= 2q\big(\log q - (\gamma_0 + \log 2\pi) - (\log 2)\iota_q(-ab^{-1})\big) \\ &\qquad + 2M_+(q;a,b) + O(\log q), \end{aligned}$$

where $M_+(q;a,b) = \sum_{\substack{d\mid q \\ d>1}} \sum_{\substack{\chi \pmod d \\ \chi \text{ primitive}}} |\chi(a)+\chi(b)|^2 \dfrac{L'(1,\chi)}{L(1,\chi)}$.

The term $(\log 2)\iota_q(-ab^{-1})$ is nonzero precisely when $a \equiv -b \pmod q$, depressing the variance in exactly the cases we wanted.

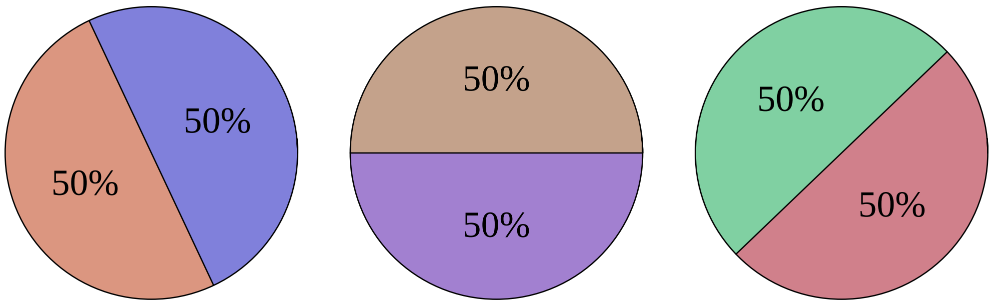
How unfair are prime number races?

- What is “Chebyshev’s bias”?
- 20th-century technology
- My current research on two-way races
  - — Limiting distributions and their variance
  - — “Unfairness measure”: dependence on modulus
  - — “Unfairness measure”: dependence on residue classes
- My current(ish) research on three-way races
- Final sound bite (two-way races)

Feuerverger and I examined various "three-way races", unearthing asymmetries that are completely unrelated to the previously known biases.

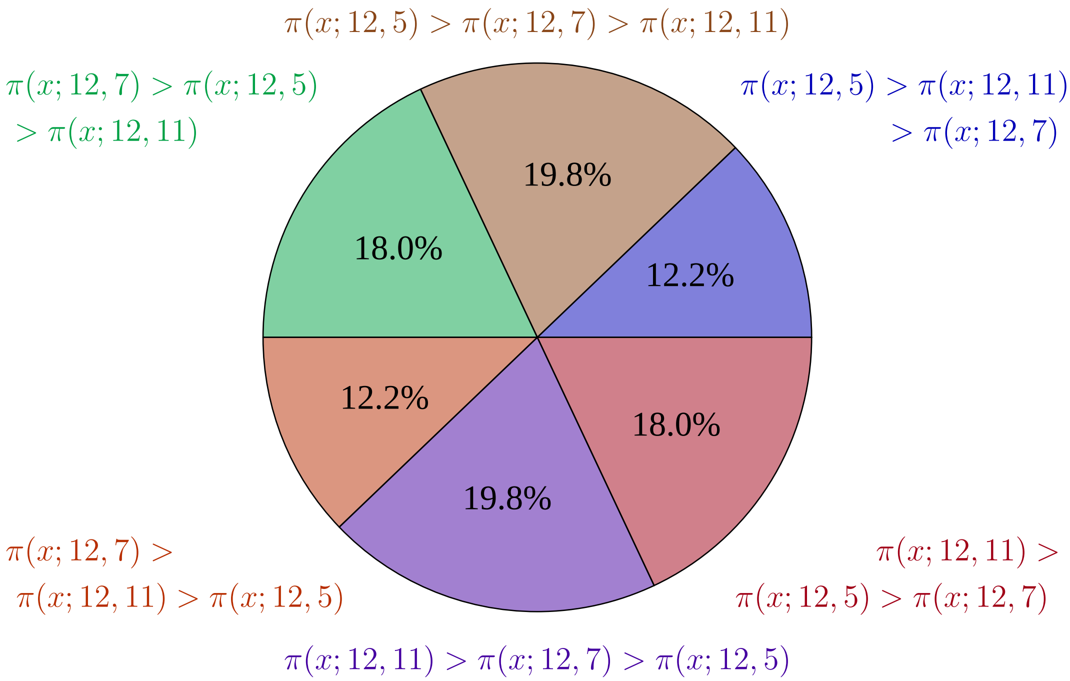
For example, 5, 7, and 11 are all nonsquares (mod 12). Therefore each two-way race is unbiased:

$\pi(x;12,5) > \pi(x;12,7)$ $\pi(x;12,5) > \pi(x;12,11)$ $\pi(x;12,7) > \pi(x;12,11)$



$\pi(x;12,7) > \pi(x;12,5)$ $\pi(x;12,11) > \pi(x;12,5)$ $\pi(x;12,11) > \pi(x;12,7)$

However, the corresponding three-way race is not:



$\pi(x;12,5) > \pi(x;12,7) > \pi(x;12,11)$
$\pi(x;12,7) > \pi(x;12,5) > \pi(x;12,11)$
$\pi(x;12,5) > \pi(x;12,11) > \pi(x;12,7)$
19.8%
18.0%
12.2%
12.2%
19.8%
18.0%
$\pi(x;12,7) > \pi(x;12,11) > \pi(x;12,5)$
$\pi(x;12,11) > \pi(x;12,5) > \pi(x;12,7)$
$\pi(x;12,11) > \pi(x;12,7) > \pi(x;12,5)$

Let $\delta_{q;a_1,a_2,a_3}$ be the "probability" that $\pi(x;q,a_1) > \pi(x;q,a_2) > \pi(x;q,a_3)$.

Theorem. Assume GRH and LI. Suppose that a_1, a_2, a_3 are distinct reduced residues $\pmod q$ such that $a_1^2 \equiv a_2^2 \equiv a_3^2 \pmod q$ and a_1, a_2, a_3 are either all squares or all nonsquares modulo q. Then

$$\delta_{q;a_1,a_2,a_3} = \frac{1}{2\pi} \tan^{-1} \frac{\sqrt{Q\big(V(q;a_1,a_2), V(q;a_2,a_3), V(q;a_1,a_3)\big)}}{W - 2V(q;a_1,a_3)} + o(1),$$

where $Q(x,y,z) = -x^2 - y^2 - z^2 + 2xy + 2yz + 2xz$ (in particular, symmetric in x, y, z) and $W = V(q;a_1,a_2) + V(q;a_2,a_3) + V(q;a_3,a_1)$.

This is largest when the variance $V(q;a_1,a_3)$ is largest, and hence the relative sizes of the $\delta_{q;a_1,a_2,a_3}$ as the three residue classes are permuted can be predicted using our knowledge of the variances described earlier.

Let $\delta_{q;a_1,a_2,a_3}$ be the "probability" that $\pi(x;q,a_1) > \pi(x;q,a_2) > \pi(x;q,a_3)$.

Theorem. Assume GRH and LI. Suppose that a_1, a_2, a_3 are distinct reduced residues $\pmod{q}$ such that $a_1^2 \equiv a_2^2 \equiv a_3^2 \pmod{q}$ and a_1, a_2, a_3 are either all squares or all nonsquares modulo q. Then

$$\delta_{q;a_1,a_2,a_3} = \frac{1}{2\pi} \tan^{-1} \frac{\sqrt{Q\big(V(q;a_1,a_2), V(q;a_2,a_3), V(q;a_1,a_3)\big)}}{W - 2V(q;a_1,a_3)} + o(1),$$

where $Q(x,y,z) = -x^2 - y^2 - z^2 + 2xy + 2yz + 2xz$ (in particular, symmetric in x, y, z) and $W = V(q;a_1,a_2) + V(q;a_2,a_3) + V(q;a_3,a_1)$.

Corollary. Under the same hypotheses,

$$\delta_{q;a_1,a_2,a_3} = \frac{1}{6} + O\left(\frac{\log\log q}{\log q}\right).$$

How unfair are prime number races?

- What is “Chebyshev’s bias”?
- 20th-century technology
- My current research on two-way races
  - — Limiting distributions and their variance
  - — “Unfairness measure”: dependence on modulus
  - — “Unfairness measure”: dependence on residue classes
- My current(ish) research on three-way races
- Final sound bite (two-way races)

The last task ahead of me in my current research on two-way races is to work out versions of the asymptotic formulas for the densities $\delta_{q;a,b}$ with explicit constants.

Once this is done, actual numerical bounds on the $\delta_{q;a,b}$ can be obtained, and a finite computation (up to $q = 1000$, perhaps) will be able to verify the following tables:

Top Ten Most Unfair Races

Modulus	Winner	Loser	Proportion
24	5	1	99.9987%
24	11	1	99.9982%
12	11	1	99.9976%
24	23	1	99.9888%
24	7	1	99.9833%
24	19	1	99.9718%
8	3	1	99.9568%
12	5	1	99.9206%
24	17	1	99.9124%
3	2	1	99.9064%

Top Ten* Most Unfair Races
(Prime Modulus Category)

Modulus	Winner/Loser**	Proportion
3	2/1	99.9064%
5	2/1	95.2035%
7	3/1	87.4035%
7	6/1	84.5209%
11	7/1	76.0558%
11	2/1	73.0668%
13	6/1	72.4288%
11	10/1	71.3943%
13	5/1	70.3981%
13	2/1	67.8108%

* There are a total of 25 distinct prime-modulus race proportions that exceed 60%.

Top Ten* Most Unfair Races
(Prime Modulus Category)

Modulus	Winner/Loser**	Proportion
3	2/1	99.9064%
5	2/1	95.2035%
7	3/1	87.4035%
7	6/1	84.5209%
11	7/1	76.0558%
11	2/1	73.0668%
13	6/1	72.4288%
11	10/1	71.3943%
13	5/1	70.3981%
13	2/1	67.8108%

** Most lines represent several races to the given modulus: for example, the third line (modulus 7) represents the races 3/1, 3/2, 5/2, 5/4, 6/2, and 6/4.

Greg Martin
University of British Columbia

“Biases in the Shanks-Rényi prime number race”, with A. Feuerverger

“Asymmetries in the Shanks-Rényi prime number race”

“Inequities in the Shanks-Rényi prime number race” (in preparation)

“Inequities in three-way prime number races” (in preparation)

`http://www.math.ubc.ca/~gerg`