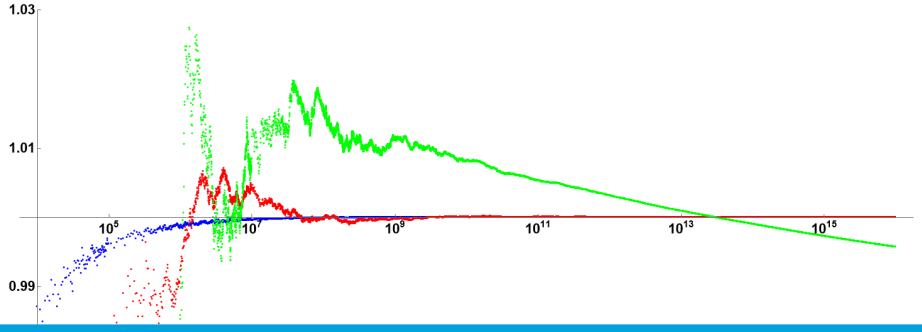


Priemgetallen-heuristiek

Benne de Weger

Vakantiecursus 2023

Priemgetallen en kansen 1/2

Wat is de kans dat een *willekeurig* getal x in de buurt van X een priemgetal is?

Strategie: tel priemgetallen in een interval $I_{X,\Delta} = (X-\Delta, X+\Delta]$ met $\Delta \ll X$:

$$P(x \in I_{X,\Delta} \text{ priem}) = \frac{\text{aantal priemgetallen in } I_{X,\Delta}}{\text{totaal aantal getallen in } I_{X,\Delta}} = \frac{\pi(X+\Delta)-\pi(X-\Delta)}{2\Delta}.$$

Priemgetalstelling: $\pi(x) \sim \dfrac{x}{\log x}$, interpreteren we 'heuristisch': $\pi(x) \approx \dfrac{x}{\log x}$:

$$P(x \in I_{X,\Delta} \text{ priem}) \approx \frac{\dfrac{X+\Delta}{\log(X+\Delta)} - \dfrac{X-\Delta}{\log(X-\Delta)}}{2\Delta}.$$

Priemgetallen en kansen 2/2

Hoe maken we chocola van $\log(X \pm \Delta)$? Gebruik de "ingenieurs-logaritme": als $\Delta \ll X$ dan is $\log(X \pm \Delta) \approx \log X \pm \frac{\Delta}{X} \approx \log X$, dus vrijwel constant:

$$P(x \in I_{X,\Delta} \text{ priem}) \approx \frac{\frac{X+\Delta}{\log X} - \frac{X-\Delta}{\log X}}{2\Delta} = \frac{(X+\Delta)-(X-\Delta)}{2\Delta \log X} = \frac{2\Delta}{2\Delta \log X},$$

dus

$$\boxed{P(x \approx X \text{ priem}) \approx \frac{1}{\log X}}.$$

Opgaven

Opgave 3.2.1 Kun je deze redenering ook maken met $\pi(x) \sim \mathrm{li}(x) = \int_2^x \frac{1}{\log t}\,\mathrm{d}t$?

[[Hint: vervang $\log t$ door een constante benadering.]]

Opgave 3.2.3 Hoeveel priemgetallen zijn er met b bits in hun binaire schrijfwijze, dus in $[2^{b-1}, 2^b - 1]$?
Hoe groot moet je b kiezen om dit aantal op 3×10^{80} te laten uitkomen? (Dit is het geschatte aantal elementaire deeltjes in het universum).

Opgave Wat is de kans dat een willekeurige *oneven* $x \approx X$ een priemgetal is?

Priemgetallen in congruentieklassen 1/2

Neem een *modulus* q en een $a \in \{0, 1, \ldots q-1\}$, met $\text{ggd}(a, q) = 1$.
$\pi(x, q, a) = \#\{p \leq x \mid p \equiv a \pmod{q}\}$ de telfunctie van priemgetallen $\equiv a \pmod{q}$.

Dirichlet–Siegel–Walfisz: $\boxed{\pi(x, q, a) \sim \frac{1}{\phi(q)} \cdot \text{li}(x)} \sim \frac{1}{\phi(q)} \cdot \frac{x}{\log x}$.

m.a.w. priemgetallen zijn *eerlijk verdeeld* over $\phi(q)$ mogelijke $a \pmod{q}$. Kansen:

$$P(x \equiv a \pmod{q} \text{ en } x \text{ is priem}) = P(x \equiv a \pmod{q}) \cdot P(x \text{ is priem}),$$

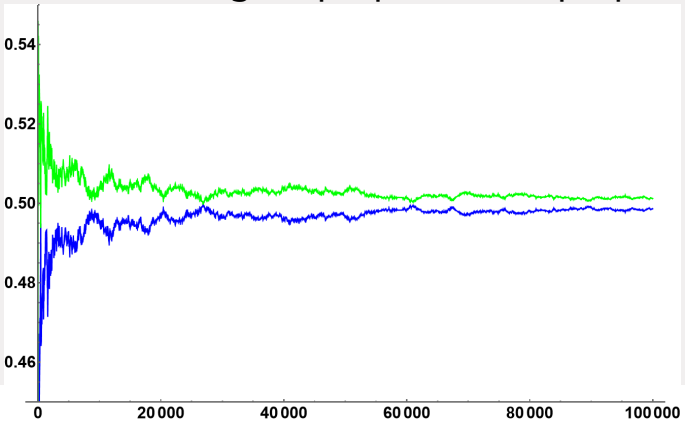
m.a.w.: kansen op gebeurtenissen $\equiv a \pmod{q}$ en is priem zijn *onafhankelijk*.

TU/e

Priemgetallen in congruentieklassen 2/2

Een experiment kan nooit kwaad.
Pari-opgave 3.3.1

```
tel(q,x) = t = vector(q-1);
           forprime(p=1, x, if(gcd(p,q)==1, t[p%q]++)); t
tel(4,10^5)
[4783, 0, 4808]
tel(4,10^6)
[39175, 0, 39322]
tel(4,10^7)
[332180, 0, 332398]
```



$3 \pmod{4}$
$1 \pmod{4}$

TU/e

Chebyshev's oneerlijkheid 1/2

Het lijkt wel of $\pi(x,4,1)$ systematisch kleiner is dan $\pi(x,4,3)$.

Rubinstein en Sarnak bewezen in 1994 (onder een Riemann-hypothese):

$$\pi(x,q,a) \sim \frac{\pi(x)}{\phi(q)}\left(1 - c(q,a)\cdot\frac{1}{\sqrt{x}} + \text{ een foutterm van dezelfde grootteorde}\right)$$

met $c(q,a) = -1 +$ aantal oplossingen van $b^2 \equiv a \pmod{q}$,
dus bijvoorbeeld $c(4,1) = 1, c(4,3) = -1$

en ook: $\pi(x,4,1) < \pi(x,4,3)$ voor ongeveer 99.6% van alle x.

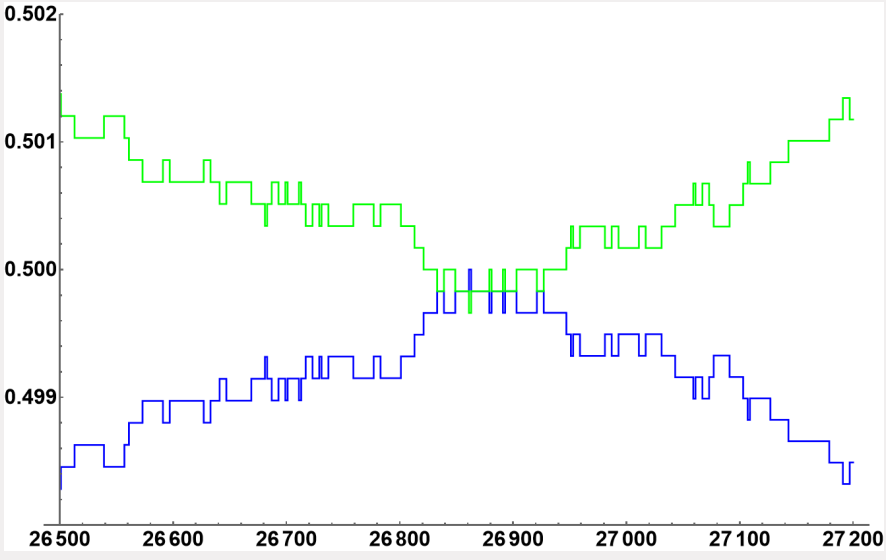
Dit is *niet* in tegenspraak met $\pi(x,4,a) \sim \frac{1}{2}\cdot \mathrm{li}(x)$ voor $a = 1,3$!

TU/e

Chebyshev's oneerlijkheid 2/2

De eerste keer dat $\pi(x,4,1) > \pi(x,4,3)$ zit bij $x = 26861$: $\longrightarrow$

Daarna gebeurt het van 616841 tot 633799 wat vaker.
Dan niet meer onder de 1 miljoen.



TU/e

Opgaven

Herinnering: $\pi(x,q,a) \approx \frac{\pi(x)}{\phi(q)}\left(1 - c(q,a)\cdot\frac{1}{\sqrt{x}}\right)$ en
$c(q,a) = -1 +$ aantal oplossingen van $b^2 \equiv a \pmod{q}$

Opgave Bereken $c(10,1), c(10,3), c(10,7), c(10,9)$, en schrijf uit wat dit betekent voor de aantallen priemgetallen die eindigen op een bepaald cijfer.

Pari-opgave Maak je resultaat uit de vorige opagve zichtbaar door $\pi(x,10,1), \pi(x,10,3), \pi(x,10,7), \pi(x,10,9)$ voor een aantal (niet te kleine) waarden van x te berekenen.

RSA-moduli 1/3

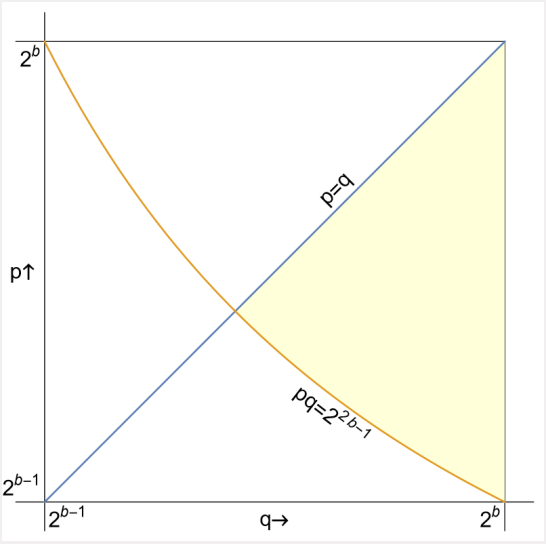
Een getal N noemen we een *goede* RSA-modulus als $N = pq$ voor priemgetallen p, q die voldoen aan:

- p en q hebben elk b bits
- N heeft $2b$ bits

voor zekere b.

Zonder verlies van algemeenheid: $p < q$.

Hoeveel goede RSA-moduli zijn er met $2b$ bits?



RSA-moduli 2/3

Precies: het aantal roosterpunten (p,q) (d.w.z. met $p,q \in \mathbb{Z}$) in het gele gebied waarvoor p en q priem zijn.

Benadering 1: $\sum \frac{1}{\log p} \cdot \frac{1}{\log q}$, waarbij de som gaat over alle roosterpunten (p,q) in het gele gebied.

Benadering 2: $\iint \frac{1}{\log p} \cdot \frac{1}{\log q}\, \mathrm{d}(p,q)$, integreer over het gele gebied.

Ondergrens-benadering 3: $\frac{1}{\log 2^b} \cdot \frac{1}{\log 2^b} \cdot V$, met $V =$ Opp(gele gebied).

Bovengrens-benadering 4: $\frac{1}{\log 2^{b-1}} \cdot \frac{1}{\log 2^{b-1}} \cdot V$.

TU/e

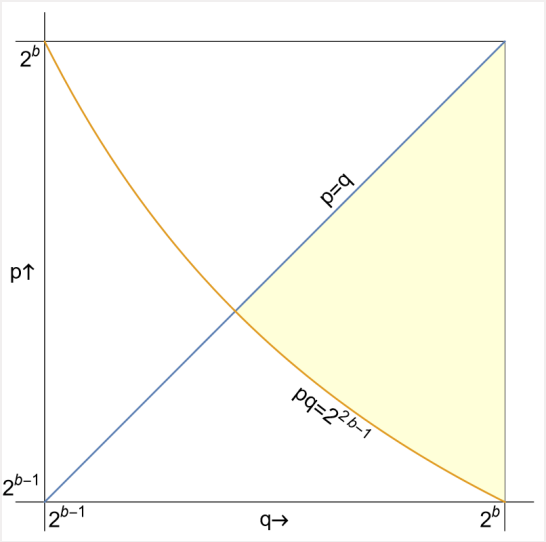
RSA-moduli 3/3

$$V = \int_{2^{b-1/2}}^{2^b} \left(q - \frac{2^{2b-1}}{q} \right) \mathrm{d}q =$$

$$\left(\frac{1}{2} q^2 - 2^{2b-1} \log q \right) \Big|_{2^{b-1/2}}^{2^b} = 2^{2b-2}(1 - \log 2).$$

Aantal goede RSA-moduli van $2b$ bits is ongeveer: $2^{2b-2}(1-\log 2) \cdot \frac{1}{b^2 (\log 2)^2} = 0.16 \frac{2^{2b}}{b^2}$ (verschil boven- en ondergrens valt binnen afrondfout in 0.16 als b groot genoeg)

Met $b = 1024$: $\approx 5 \cdot 10^{609}$.



TU/e

Hoe vind je RSA-moduli?

Met een grove methode: voor bv. 1024 bits:

- Prik 1022 random bits.
- Plak er een 1-bit achter (zodat je een oneven getal hebt).
- Plak er een 1 voor (zodat je niet minder dan 1024 bits hebt).
- Doe maximaal zes keer de Rabin-priemtest met verschillende random a's.
- Als één van de zes priemtests faalt, begin helemaal opnieuw.

Dit is efficiënt (priemtest gebruikt machtsverheffen, dat kan redelijk snel; kans op een priemgetal is ongeveer $2/\log 2^{1024} \approx \frac{1}{355}$)
en effectief (kans op een pseudopriem is astronomisch klein).

Priemtweelingen 1/4

Priemtweelingtelfunctie: $\pi_2(x) = \#\{p \leq x \mid p \text{ en } p+2 \text{ priem}\}$.

Heuristiek: $\log p$ en $\log(p+2)$ vrijwel gelijk, dus probeer als kans

$$P(p \text{ priem en } p+2 \text{ priem}) \stackrel{?}{\approx} \frac{1}{\log p} \cdot \frac{1}{\log(p+2)} \sim \frac{1}{(\log p)^2} \ldots$$

dat betekent dat $\pi_2(x) \stackrel{?}{\sim} \mathrm{li}_2(x) = \int_2^x \frac{1}{(\log t)^2}\, \mathrm{d}t \sim \frac{x}{(\log x)^2} \ldots$

Experimenteel kijken of dat een beetje klopt:

Priemtweelingen 2/4

x	$\pi_2(x)$	$\pi_2(x)/\operatorname{li}_2(x)$
10^3	35	1.06562
10^4	205	1.27805
10^5	1224	1.29672
10^6	8169	1.30806
10^7	58980	1.32546
10^8	440312	1.32016
10^9	3424506	1.32002
10^{10}	27412679	1.32038

x	$\pi_2(x)$	$\pi_2(x)/\operatorname{li}_2(x)$
10^{11}	224376048	1.32037
10^{12}	1870585220	1.32034
10^{13}	15834664872	1.32033
10^{14}	135780321665	1.32032
10^{15}	1177209242304	1.32032
10^{16}	10304195697298	1.32032
10^{17}	90948839353159	1.32032
10^{18}	808675888577436	1.32032

hier klopt duidelijk iets niet, $\pi_2(x) \sim C \cdot \operatorname{li}_2(x)$ lijkt wel goed maar niet met $C = 1$

TU/e

Priemtweelingen 3/4

Er zijn afhankelijkheden tussen de kansen dat p priem is en dat $p+2$ priem is:

als $p > 2$ priem, dan oneven, dan is $p+2$ ook oneven, dus kans maal 2

voor iedere oneven priem q is er ook zo'n correctiefactor:
dat q geen deler is van p betekent iets voor $p+2$:
$p \equiv 1, 2, 3, \ldots, q-3, q-2, q-1 \pmod{q}$ en dus
$p+2 \equiv 3, 4, 5, \ldots, q-1, 0, 1 \pmod{q}$

dus de kans dat $p+2$ geen veelvoud is van q, is niet $\frac{q-1}{q}$ maar $\frac{q-2}{q-1}$

dit geeft een correctiefactor $\frac{q-2}{q-1} / \frac{q-1}{q} = \frac{(q-2)q}{(q-1)^2} = 1 - \frac{1}{(q-1)^2}$

TU/e

Priemtweelingen 4/4

Totale correctiefactor wordt dus $2C_2$ met

$$C_2 = \prod_{q \text{ oneven priem}} \left(1 - \frac{1}{(q-1)^2}\right) = \frac{3}{4} \cdot \frac{15}{16} \cdot \frac{35}{36} \cdot \frac{99}{100} \cdot \ldots = 0.66016\ldots$$

(C_2 heet de *priemtweelingconstante*)
(zie de syllabus voor een bewijs dat oneindige product convergeert en > 0 is)

nieuwe heuristiek: $\boxed{\pi_2(x) \sim 2C_2 \cdot \mathrm{li}_2(x)} \sim 2C_2 \cdot \frac{x}{(\log x)^2}$

$2C_2 = 1.32032\ldots$ klopt heel goed met de tabel
let wel: dit is geen stelling maar een vermoeden

Constante van Brun 1/2

Frits bewees dat er oneindig veel priemgetallen zijn omdat $\sum_{p \text{ priem}} \frac{1}{p}$ divergeert.

Een heuristische interpretatie:

$$\sum_{p \text{ priem}} \frac{1}{p} = \sum_{n=1}^{\infty} \frac{1}{p_n} \approx \sum_{n=1}^{\infty} \frac{1}{n \log n} \approx \int_2^{\infty} \frac{1}{x \log x}\, \mathrm{d}x = \left.\log\log x\right|_2^{\infty} = \infty.$$

Kan dit ook voor priemtweelingen?

$$\sum_{p,p+2 \text{ priem}} \frac{1}{p} \approx 2C_2 \cdot \sum_{n=1}^{\infty} \frac{1}{n(\log n)^2} \approx 2C_2 \cdot \int_2^{\infty} \frac{1}{x(\log x)^2}\, \mathrm{d}x = -2C_2 \cdot \left.\frac{1}{\log x}\right|_2^{\infty} \text{ is eindig!}$$

Constante van Brun 1/2

Vermoeden van Brun: $\sum_{p,p+2 \text{ priemtweeling}} \left(\frac{1}{p} + \frac{1}{p+2}\right)$ convergeert, en wel naar $1.90216\ldots$.

Let op: uit het vermoedelijk eindig zijn van deze som volgt niet dat er slechts eindig veel priemtweelingen zouden zijn.

Sophie Germain-priemgetallen

Een *Sophie Germain-priemgetal* is een priemgetal p waarvoor ook $2p+1$ een priemgetal is.
Voorbeelden: $2, 3, 5, 11, 23, \ldots$.

Laat $\pi_{SG}(x)$ het aantal Sophie Germain-priemgetallen $\leq x$ zijn.

Opgave 3.5.1 Geef een argument voor de volgende heuristiek:

$$\pi_{SG}(x) \sim 2C_2 \operatorname{li}_2(x)$$

waarbij C_2 de priemtweelingconstante is.

Patronen in opeenvolgende priemgetallen 1/2

Laat q een priemgetal zijn, en $\mathbf{a} = (a_1, a_2, \ldots, a_r)$ een *patroonvector* van restklassen modulo q, met $\text{ggd}(a_i, q) = 1$.
Lemke Oliver en Soundararajan bestuderen r-tallen *opeenvolgende* priemgetallen $(p_1, p_2, \ldots, p_r)$ waarvoor $p_i \equiv a_i \pmod{q}$, en hun telfunctie

$$\pi(x, q, \mathbf{a}) = \#\{p_1 \leq x \mid p_1, \ldots, p_r \text{ opeenvolgend en } p_i \equiv a_i \pmod{q}\}.$$

A priori verwacht je

$$\pi(x, q, \mathbf{a}) \sim \frac{\text{li}(x)}{\phi(q)^r}.$$

Dat vermoeden staat wel, maar ze vinden á la Chebyshev ook oneerlijkheden, en veel grotere.

TU/e

Patronen in opeenvolgende priemgetallen 2/2

Experiment: $\pi(x_0, 10, (f, g))$ voor $f, g \in \{1, 3, 7, 9\}$ met $\pi(x_0) = 10^8 + 2$:

	$g = 1$	$g = 3$	$g = 7$	$g = 9$
$f = 1$	4623042	7429438	7504612	5442345
$f = 3$	6010982	4442562	7043695	7502896
$f = 7$	6373981	6755195	4439355	7431870
$f = 9$	7991431	6372941	6012739	4622916

(dus 7429438 is het aantal keer onder de eerste $10^8 + 2$ oneven priemgetallen dat na een priemgetal eindigend op een 1 er eentje eindigend op een 3 volgt)

Ze geven een heuristische gedetailleerde formule, zie syllabus.

TU/e

Alda-priemgetallen 1/2

Een priemgetal heet *Alda-priemgetal* als het samen met de 3 volgende priemgetallen twee priemtweelingen vormt. Voorbeelden: $5, 11, 101, \ldots$. Hun telfunctie: $\pi_{\text{🐷}}(x)$.

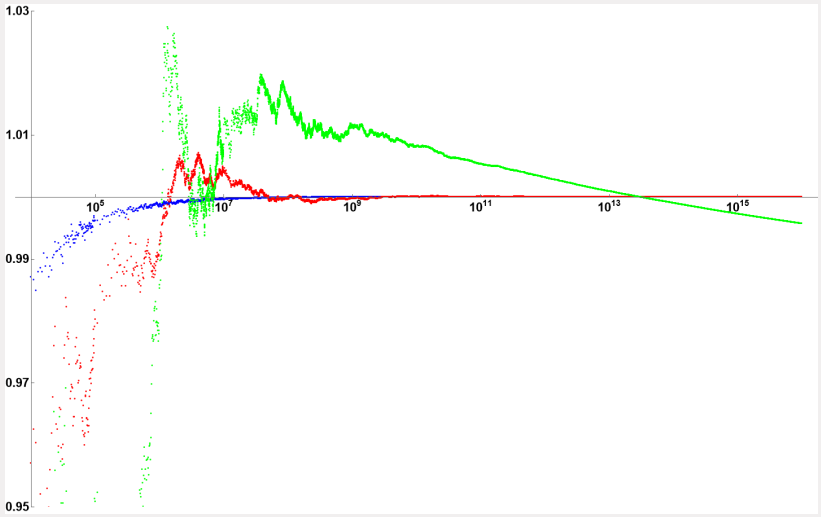
Opgave 3.7.1 Verklaar de door Efthymios Sofos gevonden heuristiek

$$\pi_{\text{🐷}}(x) = 4C_2^2 \cdot \mathrm{li}_3(x) \quad \text{waarbij } \mathrm{li}_3(x) = \int_2^x \frac{1}{(\log t)^3}\,\mathrm{d}t \sim \frac{x}{(\log x)^3}.$$

Experimenten laten hier nog veel gekkere afwijkingen zien van de heuristiek.

TU/e

Alda-priemgetallen 2/2



priems: $\dfrac{\pi(x)}{\mathrm{li}(x)}$

priemtweelingen: $\dfrac{\pi_2(x)}{2C_2 \cdot \mathrm{li}_2(x)}$

Alda-priems: $\dfrac{\pi_{\text{🐷}}(x)}{4C_2^2 \cdot \mathrm{li}_3(x)}$

We hebben geen idee wat er hier aan de hand is.

TU/e