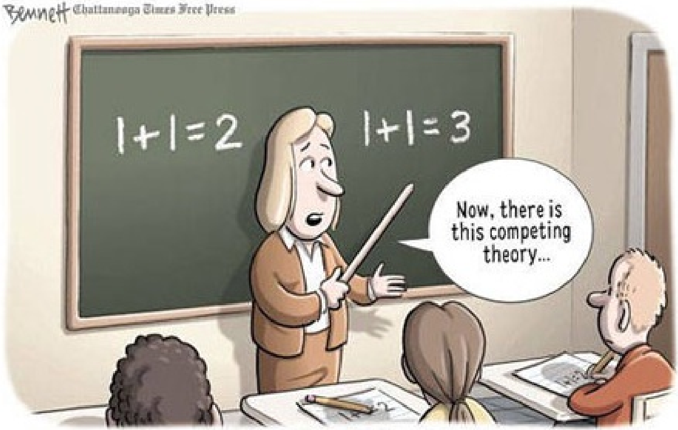
Arithmetic biases

Ghaith Hiary

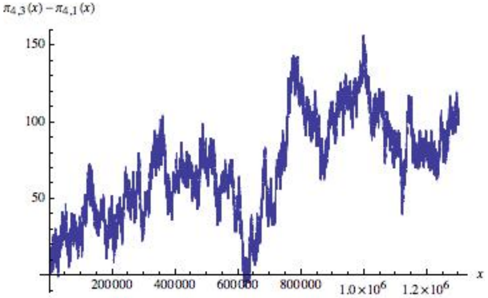

Bennett Chattanooga Times Free Press
1+1=2
1+1=3
Now, there is this competing theory...

Many important mathematical objects are quite complicated.

But they still can be analyzed by modeling them using random variables that are fair.

"Bias" occurs when it is discovered that actual behavior does not conform to our fair random model.

An example of this is Chebychev's bias. This is related to the prime numbers.



The prime numbers are those numbers greater than 1 that aren't divisible by numbers smaller than themselves and greater than 1.

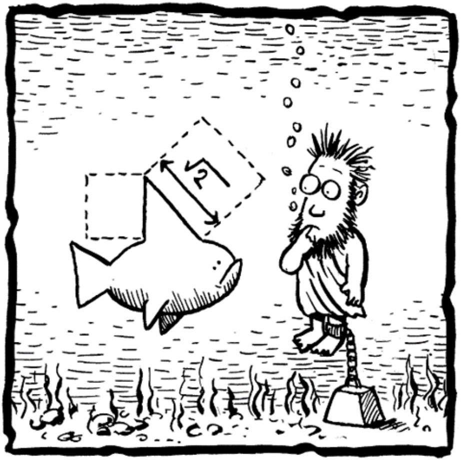
1	**2**	**3**	4	**5**
6	**7**	8	9	10
11	12	**13**	14	15
16	**17**	18	**19**	20
21	22	**23**	24	25

• **Prime** • Composite

Every positive number can be expressed in just one way as the product of prime numbers; e.g. 12 is made up of two 2s and one 3.

This is the fundemanetal theorem of arithmetic. It is a nontrivial fact!

It follows from the fundamental theorem of arithmetic that $\sqrt{2}$ cannot be written in the form p/q, where p and q are integers.

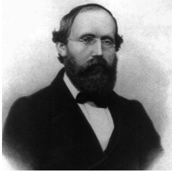


Much progress has been made on the study of the primes

Euclid proved that there are infinitely many primes

Gauss and others studied the prime counting function

$$\pi(x) = \text{number of primes} \leq x$$



Riemann studied the function

$$\zeta(s) = \prod_{\text{primes } p} \frac{1}{1 - p^{-s}} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

But *exactly* how many primes are there?

Set	Number of members $\leq$ 10,000
Powers of 2	14
Even numbers	5,000
Primes	1,229
Primes $1 + 4k$	609
Primes $3 + 4k$	619

There are infinitely many powers of 2, but among the first 10,000 numbers, there are only 14 powers of 2: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192.

In general, among the first X numbers, there are about $\log_2 X$ powers of 2. This is much less than the number of even numbers in that range.

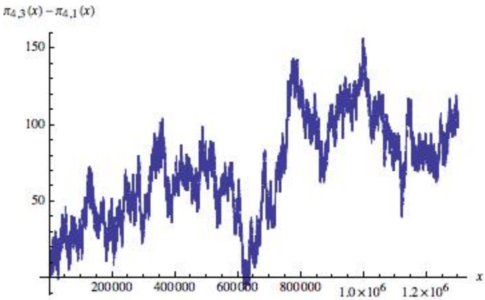
$$\lim_{x\to\infty} \frac{\pi(x)}{\mathrm{Li}(x)} = 1 \text{ where } \mathrm{Li}(x) = \int_2^x \frac{1}{\ln t}\, dt \sim \frac{x}{\ln x}$$

$x = 8^j$	$\pi(x)$	$\mathrm{Li}(x) - \pi(x)$
3	97	6
4	564	12
5	3512	31
6	23000	68
7	155611	128
8	1077871	350
9	7603553	829
10	54400028	1447
11	393615806	3585
12	2874398515	13543
13	21151907950	25739
14	156661034233	59034
15	1166746786182	262853

Notice that the number of primes $\leq$ 10,000 of the form $3 + 4k$ is more than those of the form $1 + 4k$.

Set	Number of members $\leq$ 10,000
Primes $1 + 4k$	609
Primes $3 + 4k$	619

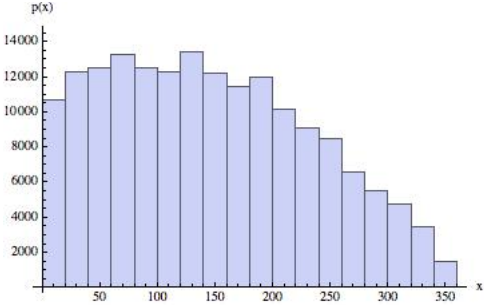
This is not a coincidence; this phenomenon seems to persist for a long time. It is an example of an arithmetic bias!



(Yitang Zhang, 2013) There are infinitely many pairs of primes differing by at most 70 million.

Is this what one expects if the primes behaved like "fair" random variables?

Here is a histogram of prime gaps for 20,000 primes around 10^{13}. Is this what we expect if we assume "random behavior"?



References:

- Jordan Ellenberg, The beauty of bounded gaps, Slate Magazine, 2013.
- Michael Rubinstein and Peter Sarnak, Chebyshev's bias, Experiment. Math. Volume 3, Issue 3 (1994), 173-197.