Benford's **Law** and irrational rotation

subtitle

How frequently do powers of 2 start with digit 7 ?

A: Benford's distribution (holds for a^n for all $a \neq 10^k$

First publication in 1880 by Simon Newcomb

Benford's distribution: 1937 "law of anomalous numbers"

digit	1	2	3	4	5	6	7	8	9
frequency	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6

%

0.301 0.176 0.125 ...

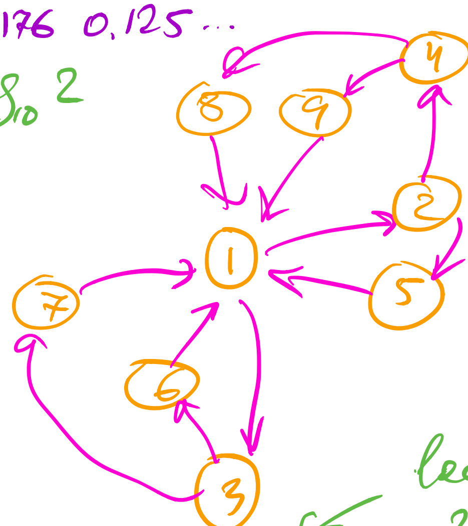
$= \log_{10} 2$

Nathan's idea:

$p_1 = p_5 + p_6 + p_7 + p_8 + p_9$

$p_2 = p_4 + p_5$

$p_3 = p_6 + p_7$



leading dig of 2^a begins w d

$$p_d = \log_{10}\left(1 + \frac{1}{d}\right), \text{ so } = \log_{10}\frac{d+1}{d} = \log_{10}(d+1) - \log_{10} d$$

base 10 world

$p_1 = \log_{10}(1+1) = \log_{10} 2$

$p_2 = \log_{10}\left(1+\frac{1}{2}\right) =$

why is this so ?

What about combinations of digits?

Say, can 2^n begin with 2021?

or with 1234567?

Yes, just switch to base

$b = 10^4$ for 2021

or $b = 10^8$ for 12...7 and

so probabilities are $\log_{10^4} \frac{2022}{2021}$

and ...

But, why stop here?

First digits

Fib. numbers, Lucas #s

Factorials, $n!$

Binomials $\binom{n}{k}$

also satisfy ~~Benford~~'s distribution

2^n $n=1, .., 20$

2	4	8	16	32	64	128	256	512	1024
2048	4096	8192	16384	32768	65536	131072	262144	524288	1048576

N	d	1	2	3	4	5	6	7	8	9
10		3	2	1	1	1	1	0	1	0
20		6	4	2	2	2	2	0	2	0
30		9	6	3	3	3	3	0	3	0
40		12	8	4	4	4	4	0	4	0
50		15	10	5	5	5	4	1	5	0
60		18	12	6	6	6	4	2	5	1
100		30	17	13	10	7	7	6	5	5

$$2^{46} = 70368744177664$$
$$2^{53} = 9007199254740992$$

- So from the first 100, 1000, 10000, 10^5, and 10^6 we expect:

Theory (Benford)

	1	2	3	4	5	6	7	8	9
100	30	18	12	10	8	7	6	5	5
1000	301	176	125	97	79	67	58	51	46
10^4	3010	1761	1249	969	792	669	580	512	458
10^5	30103	17609	12494	9691	7918	6695	5799	5115	4576
10^6	301030	176091	124939	96910	79181	66947	57992	51153	45757

Real statistics (for d=2):

	1	2	3	4	5	6	7	8	9
100	30	17	13	10	7	7	6	5	5
	301	176	125	97	79	69	56	52	45
	3010	1761	1249	970	791	670	579	512	458
10^6	301030	176093	124937	96911	79182	66947	57990	51154	45756

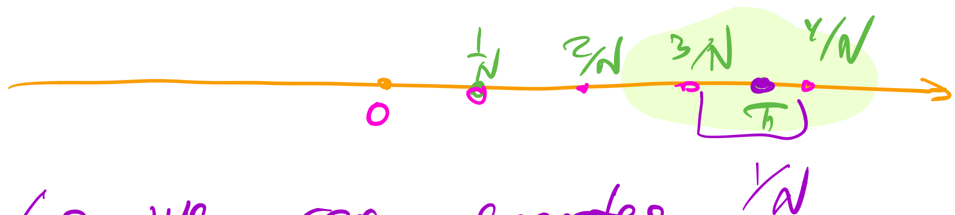
for $N = 10^6$

	Exp. Actual	Theory Benford
1	301030	301030
2	176093	176091
3	124937	124939
4	96911	96910
5	79182	79181
6	66947	66948
7	57990	57992
8	51154	51153
9	45756	45757

Rational approximations

If α ($= \sqrt{2}$ or π) is an irrational number, how well can it be approximated by rationals?
Say, if we want the denominator q of the rational number be at most N?

$$\pi = \frac{p}{q}, \quad q \le 10$$



So we can guarantee the error (precision) of $\le \frac{1}{2N}$

E.g. for $\alpha = \pi$ (3.14) $\quad \frac{16}{5} = 3.2$

we get $0.06 \le \frac{1}{10}$

but sometimes we get lucky and the precision is much better than expected

e.g. $\frac{22}{7} - \pi = 0.00126$ ← 56 times less

and $\frac{1}{2\cdot 7} = \frac{1}{14} = 0.0714$

In fact, we can get much better!

Thm For every N (whole number) there is $q \le N$ and p (whole) s.t.

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{qN}$$

("Dirichlet Approximation Thm")

Corollary If α is irrational, then there are infin many values of q s.that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$$

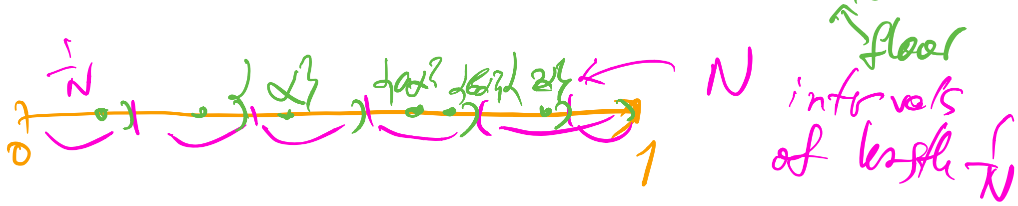
Pf of Thm Mult. by q:

$$|\alpha q - p| < \frac{1}{N}$$

i.e. the fractional part of αq is $< \frac{1}{N}$

$$= \{\alpha q\} = \alpha q - \lfloor \alpha q \rfloor$$

floor



Now take $0, \{\alpha\}, \{2\alpha\}, \dots, \{N\alpha\}$ in $[0,1)$

$$\underbrace{0, \{\alpha\}, \{2\alpha\}, \dots, \{N\alpha\}}_{N+1 \text{ numbers}}$$

so two of $\{a\alpha\}$ and $\{b\alpha\}$ will land in the same interval

(by pigeonhole principle)
(aka Dirichlet's Drawer principle)
(Box)

and so $|\{a\alpha\} - \{b\alpha\}| < \frac{1}{N}$

$$\| \quad |(a\alpha - \lfloor a\alpha \rfloor) - (b\alpha - \lfloor b\alpha \rfloor)|$$

$$= |\underbrace{(a-b)}_{q}\alpha - \underbrace{\lfloor a\alpha \rfloor + \lfloor b\alpha \rfloor}_{p}| < \frac{1}{N}$$

$a \neq b$ and $a, b \leq N$
so $|a-b| \leq N$

Pf of the corollary:

Take $N = 2$ (e.g.)
Find $q \leq N$ s.t. $|\alpha - \frac{p}{q}| < \frac{1}{qN}$
(by Thm)

so $\leq \frac{1}{q^2}$

Let $\varepsilon = |\alpha - \frac{p}{q}| > 0$, and take

$\frac{1}{N_2} < \varepsilon$ $\quad N_2 \geq \frac{1}{\varepsilon}$, then find q_2, p_2

s.t. $\left|\alpha - \frac{p_2}{q_2}\right| \le \frac{1}{q_2 N_2} < \varepsilon$, this $q_2 \ne q$

; and so on, take $\varepsilon_2 = \left|\alpha - \frac{p_2}{q_2}\right| > 0$

and $N_3 > \frac{1}{\varepsilon_2}$ etc, This

gives ∞ many rational approx.

to α st $\boxed{\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}}$ "good approx"

Applications ① An irrationality test

If $\alpha = \frac{m}{n}$ is rational, and $\frac{p}{q} \ne \frac{m}{n}$

then $\left|\alpha - \frac{p}{q}\right| = \left|\frac{m}{n} - \frac{p}{q}\right| = \frac{|mq - np|}{nq} \geq \frac{1}{nq}$ ($\ne 0$, ≥ 1)

So we cannot have ∞ many "good" approximations to α

Therefore, if α has ∞-many good approximations, then α is irrational.

Ex $e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots + \frac{1}{n!} + \dots$

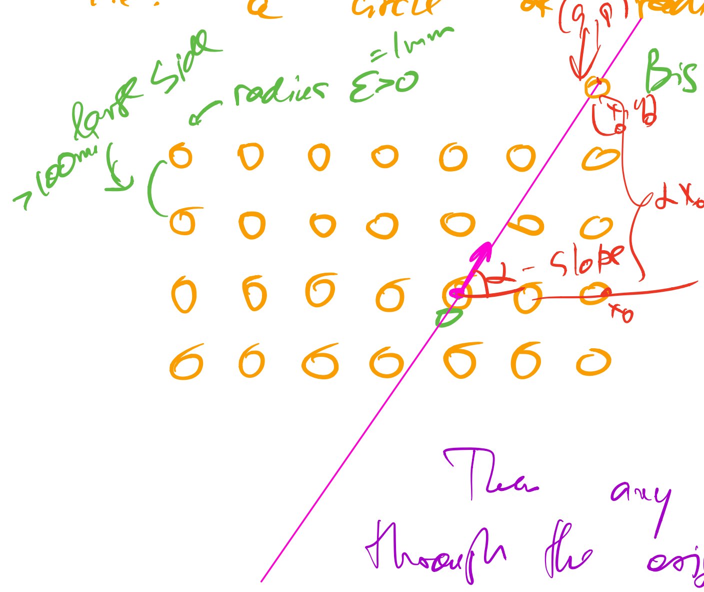
gives ∞-many good rational appr

so e is irrational

② "Application" 2

If at every point on the integer lattice we plant an arbitrary thin tree

i.e. a circle of (q,p) radius $\varepsilon > 0$



B is thin-tree forest

αx_0 Hit at (x_0, y_0) tree means that $y = \alpha x_0$ is ε-close to y_0 i.e.

$$|\alpha x_0 - y_0| < \varepsilon <$$

i.e. $\left|\alpha - \frac{y_0}{x_0}\right| < \varepsilon / x_0$

Then any straight line through the origin will intersect ∞-many trees

Now take $N > \frac{1}{\varepsilon}$ i.e. $\frac{1}{N} < \varepsilon$

and by ①. Then we have p, q , $q \le N$

$$\left|\alpha - \frac{p}{q}\right| \le \frac{1}{qN} = \frac{1}{N} \cdot \frac{1}{q} < \frac{\varepsilon}{q}$$

or $|q\alpha - p| < \varepsilon$, i.e. we'll hit the tree at pt (q,p)