The dice stuff

The Q

The original question presented was: How many times do I need to roll a 6-sided dice before there is a 50% chance I'll get ten sixes in a row?

That's a hard question to answer… so first we find out how to calculate probability:

$$P = \frac{N_{vc}}{T_c}$$

Simply put, the probability of getting so many same numbers in a row is the number of configurations that have the sequence (valid combinations) divided by the total number of configurations.

So, for a coin (a two-sided dice) being flipped 4 times, the chances of seeing 3 heads in a row are 0.1875 (18.75%)

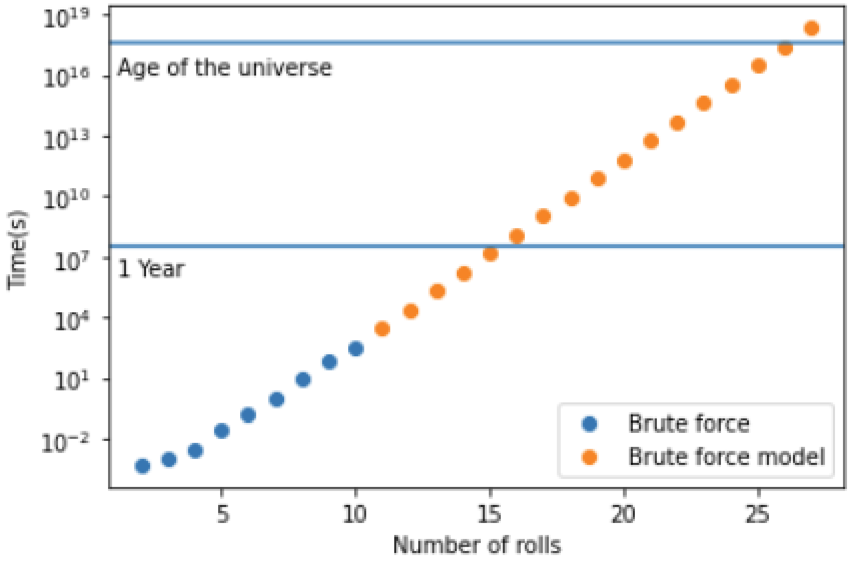
$$\left.\begin{array}{l}
(H,H,H,H)\\
(H,H,H,T)\\
(H,H,T,T)\\
(H,H,T,H)\\
(H,T,T,H)\\
(H,T,H,H)\\
(H,T,H,T)\\
(H,T,T,T)\\
(T,H,H,H)\\
(T,H,H,T)\\
(T,H,T,H)\\
(T,H,T,T)\\
(T,T,H,H)\\
(T,T,H,T)\\
(T,T,T,H)\\
(T,T,T,T)
\end{array}\right\} = \frac{3}{16} = 0.1875$$

Finding Probability

$$P = \frac{N_{vc}}{T_c} = \frac{f(F, n_r, n_c)}{F^{n_r}}$$

$P = Probabilty\ of\ finding\ a\ valid\ combination$
$T_c = Total\ combinations$
F = Number of faces on the Dice
n_r = Number of rolls
n_c= Number of successive same rolls in a row (combo)

Need a function to find N_{vc}, number of valid combinations



If we try and find N_{vc} manually we quickly run into the problem that it will take an obnoxiously long time. For reference it would take the age of the universe to manually find N_{vc} for 26 rolls of a 6 sided dice… we will be looking at millions of rolls!

More Nomenclature

{*x*,*x*,*x*,*x*} This indicates 4 rolls (i.e., $n_r = 4$)

Each *x* can take any value [1,2,3…F]

$F \geq 2$

n_c is the number of same rolls in a row (combo)

$n_c \geq 2$

$i = n_r - n_c$ This can be described as iteration number

$n_r \geq n_c$ (you can't have a combo higher than number of rolls)

$N_{vc}[i]$ *is the number of valid combinations of iteration* i

When looking for a combination of a specific number:
$N_{vc}[0] = 1$

When looking for a combination of any number:
$N_{vc}[0] = F$

Any negative integer of i, returns 0
$N_{vc}[\mathbb{Z}^{-}] = 0$

Specific number Combinations (rolling 1s in a row)

Adding one roll at a time

F=3, n_c=3, n_r=3, i = 0

{x,x,x} only one valid combination [1,1,1] (i.e., $N_{vc}[0] = 1$)

F=3, n_c=3, n_r=4, i = 1

{x,x,x,x}

{x,x,x} + {x} all the combinations from the previous step with each face on the dice

{x,1,1,1} every valid combination by adding the new dice

{1,1,1,1} the overlap of the two groups

These two sets necessarily contain all valid combinations

F=3, n_c=3, n_r=5, i = 2

{x,x,x,x,x}

{x,x,x,x} + {x} all the combinations from the previous step with each face on the dice

{x,x,1,1,1} every valid combination by adding the new dice

{x,1,1,1,1} the overlap of the two groups

Adding one roll at a time

F=3, n_c=3, n_r=6, i = 3

{*x*,*x*,*x*,*x*,*x*,*x*}

{*x*,*x*,*x*,*x*,*x*} + {*x*} all the combinations from the previous step with each face on the dice $N_{vc}[i-1] \cdot F$

{*x*,*x*,*x*,*1*,*1*,*1*} every valid combination by adding the new dice F^{i}

{x,x,*1*,*1*,*1*,*1*} the overlap of the two groups F^{i-1}

$$N[i] = N[i-1] \cdot F + F^{i} - F^{i-1}$$

We're not done yet! This may look like a working solution, but something nefarious happens when i becomes greater than n_c

Adding one roll at a time

$F=3$, $n_c=3$, $n_r=7$, $i = 4$

$\{x,x,x,x,x,x,x\}$

{x,*x,x,x,x,x*} + {*x*} all the combinations from the previous step with each face on the dice

{x,*x,x,x,1,1,1*} every valid combination by adding the new dice

{x,x,x,*1,1,1,1*} the overlap of the two groups **This does not capture valid overlaps where the i^{th} term does not equal 1**

{*x,x,x*} + {*y*} the remaining overlap *y* = [2,3,4…F]

$F=3$, $n_c=3$, $n_r=8$, $i = 5$

$\{x,x,x,x,x,x,x,x\}$

{*x,x,x,x,x,x,x*} + {*x*} all the combinations from the previous step with each face on the dice

{x,*x,x,x,x,x,1,1,1*} every valid combination by adding the new dice

{x,x,x,x,x,*1,1,1,1*} the overlap of the two groups **This does not capture valid overlaps where the i^{th} term does not equal 1**

{*x,x,x,x*} + {*y*} the remaining overlap *y* = [2,3,4…F]

Adding one roll at a time

F=3, n_c=3, n_r=9, i = 6

{*x,x,x,x,x,x,x,x,x*}

{*x,x,x,x,x,x,x,x*} + {*x*} all the combinations from the previous step with each face on the dice $N_{vc}[i-1]\cdot F$

{*x,x,x,x,x,x,1,1,1*} every valid combination by adding the new dice F^{i}

{x,x,x,x,x,*1,1,1,1*} the overlap of the two groups F^{i-1}

{*x,x,x,x,x*} + {*y*} the remaining overlap *y* = [2,3,4…F] $N_{vc}[i-n_c-1]\cdot(F-1)$

$$\boxed{N_{vc}[i]=N_{vc}[i-1]\cdot F+F^{i}-F^{i-1}-N_{vc}[i-n_c-1]\cdot(F-1)}$$

Worked Example for small i

F=3

n_c=2

n_r=4

$i = n_r - n_c = 2$

$$N[i] = \boxed{N[i-1]\cdot F} + \boxed{F^i} - \boxed{F^{i-1}}$$

Subtract the overlap (purple)

$N[i-1]\cdot F = 15$

(1, 1, 1, 2),(1, 1, 1, 3),(1, 1, 2, 1),(1, 1, 2, 2),(1, 1, 2, 3),
(1, 1, 3, 1),(1, 1, 3, 2),(1, 1, 3, 3),(2, 1, 1, 2),(2, 1, 1, 3),
(3, 1, 1, 2),(3, 1, 1, 3),(1, 1, 1, 1),(2, 1, 1, 1),(3, 1, 1, 1)

12 + 3

F^i = 9

(1, 2, 1, 1),(1, 3, 1, 1),(2, 2, 1, 1),(2, 3, 1, 1),(3, 2, 1, 1),
(3, 3, 1, 1), (1, 1, 1, 1),(2, 1, 1, 1),(3, 1, 1, 1)

6 + 3

F^{i-1} = 3

(1, 1, 1, 1),(2, 1, 1, 1),(3, 1, 1, 1)

Worked Example for i greater than n_c

$F=3$

$n_c=2$

$n_r=5$

$i = n_r - n_c = 3$

$$N[i] = \boxed{N[i-1]\cdot F} + \boxed{F^{i}} - \boxed{F^{i-1}} - \boxed{N[i-n_c-1]\cdot(F-1)}$$

$N[i-1]\cdot F = 63$

(1, 1, 1, 2, 1),(1, 1, 1, 2, 2),(1, 1, 1, 2, 3),(1, 1, 1, 3, 1),(1, 1, 1, 3, 2), (1, 1, 1, 3, 3),(1, 1, 2, 1, 2),(1, 1, 2, 1, 3),(1, 1, 2, 2, 1),(1, 1, 2, 2, 2), (1, 1, 2, 2, 3),(1, 1, 2, 3, 1),(1, 1, 2, 3, 2),(1, 1, 2, 3, 3),(1, 1, 3, 1, 2), (1, 1, 3, 1, 3),(1, 1, 3, 2, 1),(1, 1, 3, 2, 2),(1, 1, 3, 2, 3),(1, 1, 3, 3, 1), (1, 1, 3, 3, 2),(1, 1, 3, 3, 3),(1, 2, 1, 1, 2),(1, 2, 1, 1, 3),(1, 3, 1, 1, 2), (1, 3, 1, 1, 3),(2, 1, 1, 1, 2),(2, 1, 1, 1, 3), (2, 1, 1, 2, 1),(2, 1, 1, 2, 2), (2, 1, 1, 2, 3), (2, 1, 1, 3, 1),(2, 1, 1, 3, 2),(2, 1, 1, 3, 3),(2, 2, 1, 1, 2), (2, 2, 1, 1, 3),(2, 3, 1, 1, 2),(2, 3, 1, 1, 3),(3, 1, 1, 1, 2),(3, 1, 1, 1, 3), (3, 1, 1, 2, 1),(3, 1, 1, 2, 2),(3, 1, 1, 2, 3),(3, 1, 1, 3, 1),(3, 1, 1, 3, 2), (3, 1, 1, 3, 3),(3, 2, 1, 1, 2),(3, 2, 1, 1, 3),(3, 3, 1, 1, 2),(3, 3, 1, 1, 3), (1, 1, 1, 1, 2),(1, 1, 1, 1, 3), (1, 1, 1, 1, 1),(1, 1, 2, 1, 1),(1, 1, 3, 1, 1), (1, 2, 1, 1, 1),(1, 3, 1, 1, 1),(2, 1, 1, 1, 1),(2, 2, 1, 1, 1),(2, 3, 1, 1, 1), (3, 1, 1, 1, 1),(3, 2, 1, 1, 1),(3, 3, 1, 1, 1)

52 + 11

F^{i} = 27

(1, 2, 3, 1, 1),(1, 3, 2, 1, 1),(1, 3, 3, 1, 1),(2, 1, 2, 1, 1),(2, 1, 3, 1, 1), (2, 2, 2, 1, 1),(2, 2, 3, 1, 1),(2, 3, 2, 1, 1),(2, 3, 3, 1, 1),(3, 1, 2, 1, 1), (3, 1, 3, 1, 1),(3, 2, 2, 1, 1),(3, 2, 3, 1, 1),(3, 3, 2, 1, 1),(3, 3, 3, 1, 1), (1, 2, 2, 1, 1),(1, 1, 1, 1, 1),(1, 1, 2, 1, 1),(1, 1, 3, 1, 1),(1, 2, 1, 1, 1), (1, 3, 1, 1, 1),(2, 1, 1, 1, 1),(2, 2, 1, 1, 1),(2, 3, 1, 1, 1),(3, 1, 1, 1, 1), (3, 2, 1, 1, 1),(3, 3, 1, 1, 1)

16 + 11

Subtract the overlap (purple)

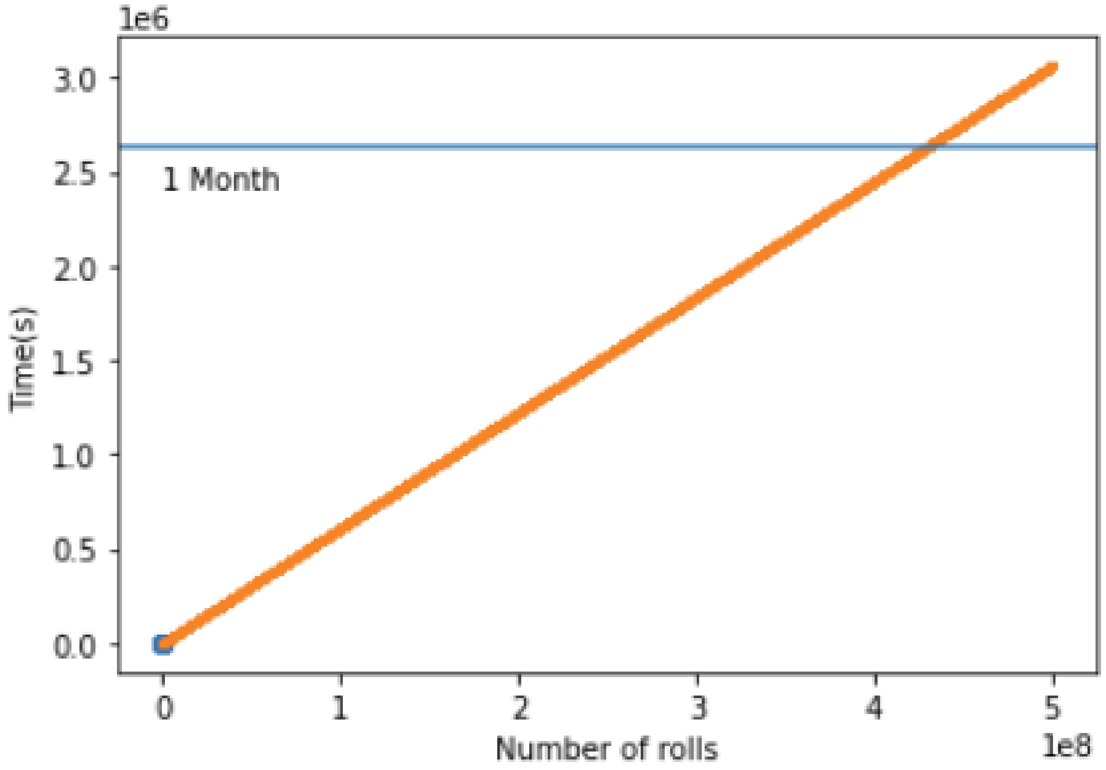
F^{i-1} = 9

(1, 1, 1, 1, 1), (1, 2, 1, 1, 1), (1, 3, 1, 1, 1),(2, 1, 1, 1, 1), (2, 2, 1, 1, 1),(2, 3, 1, 1, 1),(3, 1, 1, 1, 1),(3, 2, 1, 1, 1), (3, 3, 1, 1, 1)

$N[i-n_c-1]\cdot(F-1)$ = 2

(1, 1, 2, 1, 1),(1, 1, 3, 1, 1)

Compute time



Even though its faster than brute forcing, for large nr it’s still too slow (rolling a d6 50million times and finding combos of 10).

Tidy equation

We can rewrite slide 8 as:

$$N[i] = N[i-1]F + (F-1)\left(F^{i-1} - N[i-n_c-1]\right)$$

Assuming for large i:

$$F^{i-1} \gg N[i-n_c-1]$$

We get (which is the same as slide 6):

$$N[i] \approx N[i-1]F + (F-1)F^{i-1}$$

Let’s go a few in and simplify:

$$N[i] \approx F\left(F\left(FN[i-3] + F^{i-2} - F^{i-3}\right) + F^{i-1} - F^{i-2}\right) + F^{i} - F^{i-1}$$

$$N[i] \approx F^{3}N[i-3] + 3F^{i} - 3F^{i}$$

$$N[i] \approx F^{i}N[i-i] + i\left(F^{i} - F^{i-1}\right)$$

Since we know $N[0] = 1$:

$$N[i] \approx F^{i} + i\left(F^{i} - F^{i-1}\right)$$

Tidy equation

$$N_{vc} \approx F^{i} + i\left(F^{i} - F^{i-1}\right) \approx F^{n_r-n_c} + (n_r - n_c)(F^{n_r-n_c} - F^{n_r-n_c-1})$$

Plugging the above into equation one:

$$P \approx \frac{F^{n_r-n_c} + (n_r - n_c)(F^{n_r-n_c} - F^{n_r-n_c-1})}{F^{n_r}}$$

We get a lower limit for n_r:

$$n_r \geq \frac{PF^{n_c+1} + n_cF - n_c - F}{F - 1}$$

An upper bound can be found using the limit (as $n_r \to \infty$, $P \to 1$):

$$n_r \leq \frac{1}{1-P}\frac{PF^{n_c+1} + n_cF - n_c - F}{F - 1}$$

Solution to the original problem

F=6, n_c=10, n_r=?, P = 0.5

36279714 < n_r < 72559428

Running a simulation of n_r = 54419571 (the average of the bounds) 100 times.
52 had a combo of 10 6s in a row present.

What would take a month to calculate now takes half a second!

All Number Combinations

Adding one roll at a time

F=3, n_c=3, n_r=3, i = 0

{*x,x,x*} F valid combinations [1,1,1] [2,2,2] [3,3,3] (i.e., $N_{vc}[0] = F$)

F=3, n_c=3, n_r=4, i = 1

{*x,x,x,x*}

{*x,x,x*} + {*x*} all the combinations from the previous step with each face on the dice
{*x,1,1,1*} {x,2,2,2} {x,3,3,3} every valid combination by adding the new dice
{*1,1,1,1*} {*2,2,2,2*} {*3,3,3,3*} the overlap of the two groups

These two sets necessarily contain all valid combinations

Leaving the rest as an exercise to the reader:

$$N_{vc}[i] = N_{vc}[i-1]F + F^{i+1} - F^{i} - N_{vc}[i-n_c](F-1)$$

$$N_{vc} \approx F^{n_r-n_c+1} + (n_r - n_c)(F^{n_r-n_c+1} - F^{n_r-n_c})$$

$$\frac{PF^{n_c} + n_cF - n_c - F}{F-1} \le n_r \le \frac{1}{1-P}\frac{PF^{n_c} + n_cF - n_c - F}{F-1}$$