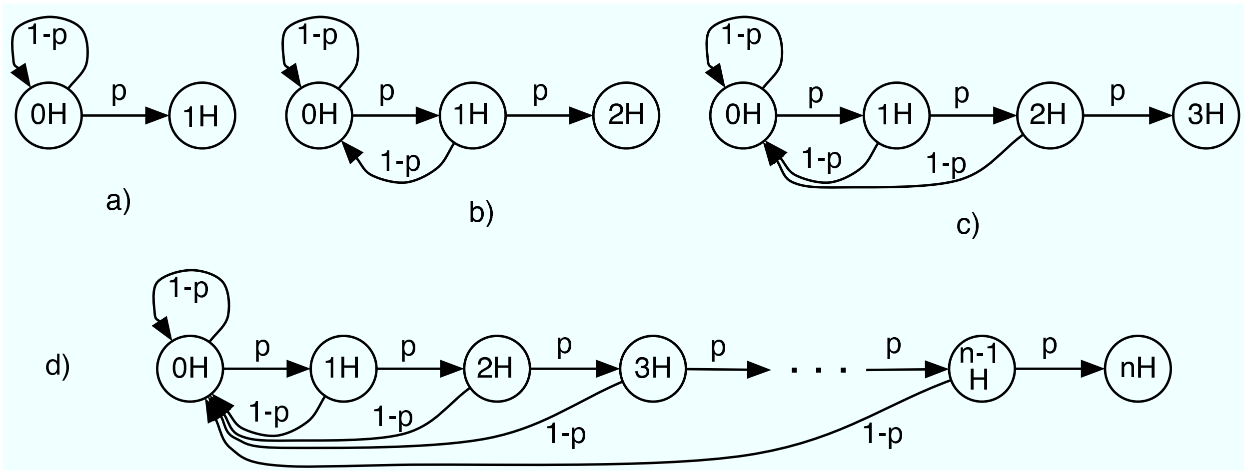
How many coin flips on average does it take to get n consecutive heads?

The process of flipping n consecutive heads can be described by a Markov chain in which the states correspond to the number of consecutive heads in a row, as depicted below. In this language, the question becomes how many steps does it take on average to get from the state 0H to the state nH?



Assume the coin has probability p of coming up heads. Begin with the case depicted in fig. (a), and let A_1 be the average number of flips on average before getting the first head. If the first flip is heads (probability p), then the answer is 1; if, on the other hand, the first flip is tails (probability $1-p$), then one flip is wasted and there remain A_1 to go. These two observations give an equation for A_1:

$$A_1 = (1-p)(1+A_1) + p \cdot 1 \ , \qquad (a1)$$

with solution

$$A_1 = \frac{1}{p} \ . \qquad (a2)$$

(This result should be familiar, since if the probability to remain in a state is $1-p$, then the average number of steps to leave the state is* $\sum_{k=1}^{\infty} k(1-p)^{k-1}p = (1/p^2)p = 1/p$.) For $p = 1/2$, we find $A_1 = 2$, so on average two flips are required to get the first head if the coin is fair.

Now consider A_2, the average number of flips to get two heads in a row (fig. (b)). Again, if the first flip is "wasted" on a tails, there's a term $(1-p)(1+A_2)$ on the right side. But now if the first flip is heads, there are two possibilities for what happens next. If the next flip is tails, the first two flips are "wasted" and we're back where we started.

* See footnote 2 below.

But if the next flip is a head, then the goal is accomplished in two flips. This gives the equation

$$A_2 = (1-p)(1+A_2) + p(1-p)(2+A_2) + p^2 \cdot 2 \ , \qquad (b1)$$

with solution

$$A_2 = \frac{1+p}{p^2} \ . \qquad (b2)$$

For $p = 1/2$, we find $A_2 = 6$, so on average six flips are required to get 2 heads in a row if the coin is fair.

Similar reasoning for A_3, the average number of flips to get three heads in a row (fig. (c)) gives

$$A_3 = (1-p)(1+A_3) + p(1-p)(2+A_3) + p^2(1-p)(3+A_3) + p^3 \cdot 3 \ , \qquad (c1)$$

with solution

$$A_3 = \frac{1+p+p^2}{p^3} \ . \qquad (c2)$$

For $p = 1/2$, we find $A_3 = 14$, so on average fourteen flips are required to get 3 heads in a row if the coin is fair.

In general, the average number of flips to get n heads in a row (fig. (d)), A_n, satisfies

$$A_n = (1-p)(1+A_n) + p(1-p)(2+A_n) + p^2(1-p)(3+A_n) + \ldots + p^{n-1}(1-p)(n+A_n) + p^n \cdot n \ . \qquad (d1)$$

Regrouping terms on the right hand side and using[1] $1 + p + p^2 + \ldots + p^{n-1} = \frac{1-p^n}{1-p}$ gives

$$\begin{aligned} A_n &= A_n(1-p)(1+p+p^2+\ldots+p^{n-1}) + (1-p)(1+2p+3p^2+\ldots+np^{n-1}) + np^n \\ &= A_n(1-p^n) + (1-p+2p-2p^2+3p^2-3p^3+\ldots+np^{n-1}-np^n) + np^n \\ &= A_n - p^n A_n + (1+p+p^2+\ldots+p^{n-1}) \ . \end{aligned}$$

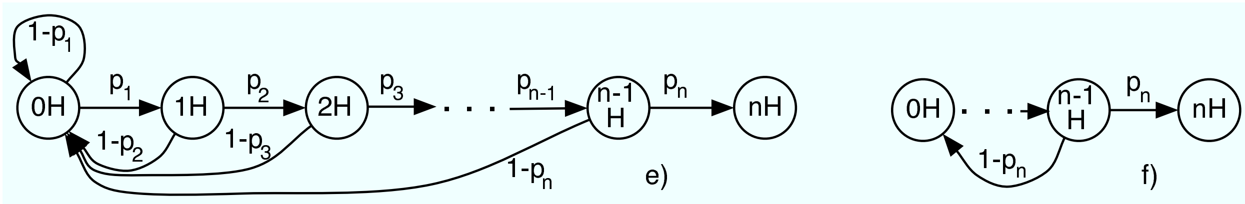
This results in

$$A_n = \frac{1+p+p^2+\ldots+p^{n-1}}{p^n} = \frac{1-p^n}{p^n(1-p)} = \frac{p^{-n}-1}{1-p} \ . \qquad (d2)$$

For $p = 1/2$, we find $A_n = 2^{n+1} - 2$ flips required to get n heads in a row if the coin is fair, and the number grows exponentially in n.

[1] To prove this, let $S_n = \sum_{k=0}^{n-1} p^k$ and note that $1 + pS_n = S_n + p^n$.

A slight generalization of this problem is to have a different probability for each successive head, i.e., to switch to a coin with probability p_j of getting a head when going for the j^{th} head in a row, as depicted in fig. (e) below:



The average number of flips A_n to get n heads in a row now satisfies

$$A_n = (1-p_1)(1+A_n) + p_1(1-p_2)(2+A_n) + p_1p_2(1-p_3)(3+A_n) + \\ \ldots + (p_1p_2\cdots p_{n-1})(1-p_n)(n+A_n) + (p_1p_2\cdots p_n)\cdot n \; . \qquad (e1)$$

Algebra similar to that leading from $(d1)$ to $(d2)$ now results in

$$A_n = \frac{1+p_1+p_1p_2+\ldots+p_1p_2\cdots p_{n-1}}{p_1p_2\cdots p_n} \; . \qquad (e2)$$

Note 1: All of the above results can derived from a single recursion equation, as suggested by fig. (f). Suppose A_{n-1}, the average number of flips required to reach $n-1$ successive heads is known. Then A_n can be determined without knowing the precise details of what happens for the first $n-1$ flips, as depicted by the ellipsis $(\cdots)$ in fig. (f). It takes an average of A_{n-1} steps to reach the state $(n-1)H$. If the next flip is heads (probability p_n), then the answer is $A_{n-1}+1$; if, on the other hand, the next flip is tails (probability $1-p_n$), then $A_{n-1}+1$ flips have been wasted and there remain A_n to go. These two observations give an equation for A_n in terms of A_{n-1}:

$$A_n = (1-p_n)(A_{n-1}+1+A_n) + p_n(A_{n-1}+1) \; , \qquad (f1)$$

with solution

$$A_n = (A_{n-1}+1)\frac{1}{p_n} \; . \qquad (f2)$$

Starting from $A_0 = 0$, the above equation gives $A_1 = 1/p_1$, $A_2 = (1/p_1+1)/p_2 = (1+p_1)/p_1p_2$, $A_3 = (1+p_1+p_1p_2)/p_1p_2p_3$, and by induction gives $(e2)$ for A_n. For $p_1 = p_2 = \ldots = p_n = p$, these are equivalent to $(a2)$, $(b2)$, $(c2)$, $(d2)$. So eq. $(f2)$, describing fig. (f), embodies the content of all of the previous equations.

Note 2: Another direct way to derive all of the above is based on the relation between figures (a) and (f). The probability to loop around k times, i.e., to get to the state $(n-1)H$ and go back to $0H$ exactly k times before ultimately getting to nH is $(1-p_n)^k p_n$, and the average number of steps for this process is $(A_{n-1}+1)k + A_{n-1} + 1 = (k+1)(A_{n-1}+1)$.

Summing over k gives $A_n = \sum_{k=0}^{\infty}(A_{n-1}+1)(k+1)(1-p_n)^k p_n$. Using[2] $1+2(1-p)+3(1-p)^2+\ldots = \frac{1}{p^2}$, it follows that

$$A_n = (A_{n-1}+1)p_n \sum_{k=0}^{\infty}(k+1)(1-p_n)^k = (A_{n-1}+1)p_n \frac{1}{p_n^2} = (A_{n-1}+1)\frac{1}{p_n} \ ,$$

reproducing the generating equation $(f2)$. (The $1/p_n$ is now recognized as the same familiar $1/p$ mentioned after eq. $(a2)$.)

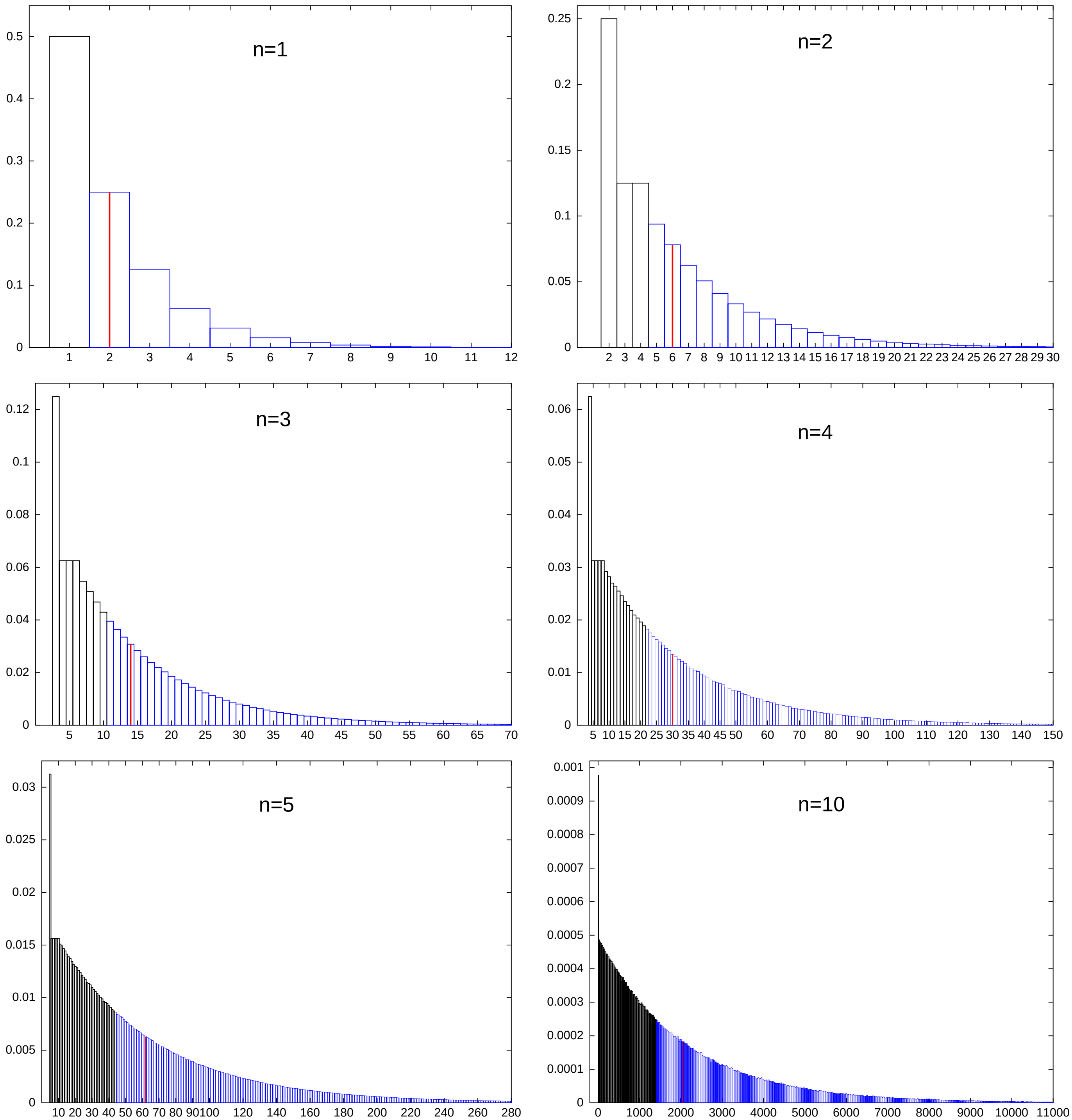
As usual, there's more structure in these probability distributions than just the average number of steps. Let $p_n(M)$ denote the probability of reaching n consecutive heads only after exactly M flips of a single coin, each flip with probability p of heads. Then the average satisfies $A_n = \sum_{M=0}^{\infty} M p_n(M)$.

For $n=1$, the probability to reach the first head in M flips is the probability of $M-1$ tails and one head, hence $p_1(M) = p^M$. The average number of flips until the first head is $\sum_{k=0}^{\infty}(k+1)(1-p)^k p = 1/p$. The probability distribution $p_1(M)$ is shown for a fair coin $(p=1/2)$ in the first figure on the next page.

Additional figures show the probability distributions for $n = 2,3,4,5,10$. In general, the probability vanishes, $p_n(M)=0$, for $M<n$ since it's impossible to have n consecutive heads with fewer than n total flips. The first non-zero probability is $p_n(M=n)=p^n$, corresponding to all heads for the first n flips. For the next n values of M, from $M=n+1$ through $M=2n$ flips, the probability is constant, $p_n(M)=p^{n+1}$, since it is fully characterized by just the last $n+1$ flips (i.e., a tail followed by n heads, and anything can happen in the first $M-(n+1)$ flips). For larger values of M, $p_n(M)$ becomes the probability of *not* having more than $n-1$ consecutive heads in the first $M-(n+1)$ flips, then followed by a final tail and n heads in the last $n+1$ flips. For example, for $M=2n+1$ flips, that probability is just all the ways *not* to have n consecutive flips in the first n flips, then the tail and n heads, so $p_n(M=2n+1)=(1-p^n)p^{n+1}$.

The probabilities $p_n(M)$ are thus related to the probability of having no more than n consecutive heads in $M-(n+1)$ flips, in turn equal to 1 minus the probability of having at least n consecutive heads in $M-(n+1)$ flips. In general, the probability of having at least n consecutive heads in N flips of a fair coin (or equivalently the probability of at least n consecutive successes in N Bernoulli trials) is difficult to write down in closed form. To provide some intution for how those numbers behave, consider the example of $N=100$.

[2] To prove this, let $S=\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$, and note that $\sum_{k=0}^{\infty} k q^{k-1} = \frac{\partial}{\partial q} S = \frac{1}{(1-q)^2}$.



Figures: The probabilities $p_n(M)$ of first flipping n consecutive heads after exactly M flips of a fair ($p = 1/2$) coin, for $n = 1, 2, 3, 4, 5, 10$. M is plotted along the horizontal axis. The red line shows the value of the average number of rolls required, eq. ($d2$): $A_n = 2^{n+1} - 2$ (resp., 2,6,14,30,2046). The regions indicated in black represent the first 50% of the probability for each of the graphs.

What is the probability $p(n)$ of n heads in a row somewhere in a sequence of 100 coin flips?

For the case of a fair coin ($p = 1/2$), the results of a numerical simulation (2 million "students" each flipping 100 times, simulated by a random number generator) are roughly given by:

n	$p(n)$	$z(n)$
2	1	16.6
3	0.9997	7.04
4	0.97	3.26
5	0.81	1.56
6	0.55	0.76
7	0.32	0.37
8	0.17	0.18
9	0.088	0.09
10	0.044	0.045
11	0.022	0.022
12	0.011	0.011
13	0.0053	0.0053
14	0.0027	0.0027
15	0.0013	0.0013
16	0.00063	0.00063

Also shown is $z(n)$, the average *number* of times that a string of n heads occurs (where by definition, for example, six consecutive heads counts as exactly two occurrences of three consecutive heads).

So if a class of 50 students were asked to generate random strings of 100 heads and tails, roughly forty of them (81%) should have at least one occurrence of five consecutive heads, roughly 27 (55%) at least six consecutive heads, and roughly 16 (32%) at least seven consecutive heads (and similarly for consecutive tails).

Since these are relatively rare events, the number of times m that a string of n consecutive heads occurs will be roughly Poisson distributed, and determined by the mean number of occurrences, $P_n(m) = \mathrm{e}^{-z(n)}\big(z(n)\big)^m/m!$. For example, since the mean number of times that five heads in a row occur in a hundred flips is $z(5) = 1.56$, the probability that five heads in a row occur m times is $P_n(m) = \mathrm{e}^{-1.56}(1.56)^m/m!$, and the probabilities of 0–5 occurrences respectively are .21,.33,.26,.13,.05,.02. Similarly, for $n = 7$, with a mean $z(7) = .37$, the probabilities for $m = 0, 1, 2, 3$ are .69,.26,.05,.01 . (For larger values of n, the average number of occurrences $z(n)$ appoaches the probability $p(n)$, so the likelihood is one occurrence if at all, and the probabilities $P_n(m)$ are dominated by $m = 0, 1$.)