Problem 1337. Off With their Heads Take n fair coins and flip them repeatedly as follows. Flip all of them. After this first flip, take all coins that show HEADs and flip them again. After the second flip, take all coins that came up HEADs and flip them again. Repeat until all HEADs are gone. Let q_n be the probability that the last group to be flipped consists of only one coin.

We have $q_0 = 0$ because there cannot be one coin. And $q_1 = 1$ because the last group to be flipped is the single coin.

(a) Is q_6 larger or smaller than q_5?

(b) Is q_{21} larger or smaller than q_{20}?

Solution. The problem was solved by Barry Cox, Joseph DeVincentis, Jim Tilley Rob Pratt, Mark Rickert, Nick Wedd, Dan Dima, Emanuele Zoccari, Jim Guilford, Richard Pulskamp, Ben Katz, and Nagendra Gulur. The values of q_n satisfy the following recurrence:

$$q_0 = 0; \quad q_1 = 1$$

$$q_n = \frac{1}{2^n - 1} \sum_{k=1}^{n-1} q_k \binom{n}{k}.$$

The recurrence holds because it is certain that at some point a TAIL will occur. When that first happens, the number of HEADs after the throw will be $k = 1$ or 2 or ... or $n - 1$. For each choice we are left with a smaller problem having probability q_k. And the probability of being in the case of k is $\frac{1}{2^n-1}\binom{n}{k}$, because there are $\binom{n}{k}$ sets of size k and the total number of possibilities excludes the case of all HEADs, and so is $2^n - 1$. This leads to a table of values that answers the question as posed.

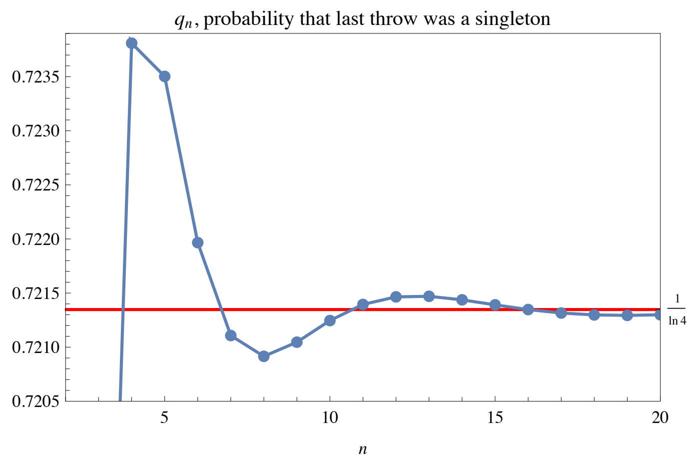
```
q[0] = 0; q[1] = 1;
q[n_] := q[n] = 1/(2^n - 1) Sum[q[k] Binomial[n, k], {k, 1, n-1}]
q /@ Range[6]
N[q /@ {5, 6, 20, 21}]
```

Out[ ]= $\left\{1, \frac{2}{3}, \frac{5}{7}, \frac{76}{105}, \frac{157}{217}, \frac{470}{651}\right\}$

Out[ ]= {0.723502, 0.721966, 0.721299, 0.721312}

These numbers are close to $\frac{1}{2\ln 2}$. The value of $2q_{200}$ is 0.693145, very close to $\ln 2$, 0.693147. This leads one to guess (incorrectly!) that the values are approaching $\frac{1}{\ln 4}$ as $n \to \infty$. Here is a plot of q_n for n up to 20. There appears to be damped oscillation about $\frac{1}{\ln 4}$. But don't be fooled! The reason I posed this problem was because of the surprising result that $\lim q_n$ does not exist.

Out[ ]=



Proof that $\lim_{n\to\infty} q_n$ does not exist. This counterintuitive result is due to Lennart Räde [2]. The solution in [2] is by O. P. Lossers, a group of solvers from the Netherlands. A later paper by Calkin, Canfield, and Wilf [1] gave three proofs that the limit does not exist. The notes here focus on the ideas in [2]. But see the third proof in [1] for a different approach using the same basic idea.

The recurrence is a natural way to express q_n, but there is another way. For the event to occur, there must be k such that one coin (there are n choices) first shows TAILs on the kth throw (this will be the last throw; the probability of this is $\left(\frac{1}{2}\right)^k$) while each other coin first showed TAIL at throw i, with $1 \le i \le k-1$. This gives

$$q_n = n \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \left(\sum_{i=1}^{k-1} \left(\frac{1}{2}\right)^i\right)^{n-1}.$$

Evaluate the inner sum and translate the index k down by 1:

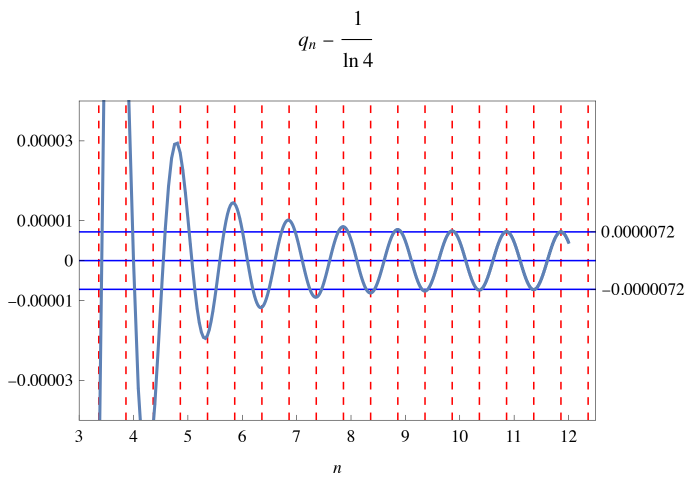
$$(1) \quad q_n = n \sum_{k=0}^{\infty} 2^{-k-1} \left(1 - 2^{-k}\right)^{n-1}.$$

(PoW reader Zoccari used this formula to get the values requested.)

A quick way to approximate (1) is by an integral. That turns out to give a value that is independent of n:

$$n \int_0^{\infty} 2^{-k-1} \left(1 - 2^{-k}\right)^{n-1} dk = \frac{1}{2 \ln 2} = 0.721347\ldots .$$

This is because the indefinite integral is $\frac{1}{\ln 4}\left(1 - 2^{-k}\right)^n$. But the independence of n is a mirage, and it fails in (1). The next plot shows the differences $q_n - \frac{1}{\ln 4}$ as n goes to 2^{12}; we see that the amplitudes are apparently not decreasing to 0. This plot uses $\log_2 n$ for the n-axis.



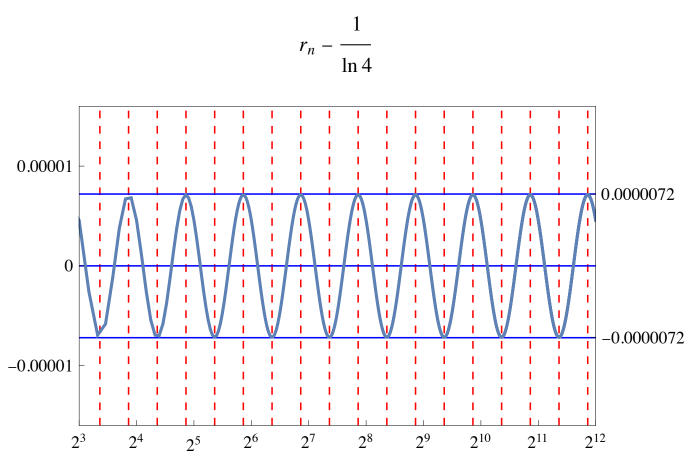
To understand this, we first use some algebraic transformations from O. P. Lossers [2]. Let $k_n = \lfloor \log_2 n \rfloor$ and $\theta_n = \operatorname{frac}(\log_2 n) = (\log_2 n) - k_n$ and $\lambda_n = 2^{\theta_n} = 2^{\operatorname{frac}(\log_2 n)}$. Then

$$q_n = \frac{n}{2} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-1} = \frac{1}{2} \sum_{j=-k_n}^{\infty} \lambda_n \left(\frac{1}{2}\right)^{j+k_n} \left(1 - \left(\frac{1}{2}\right)^{j+k_n}\right)^{n-1}$$

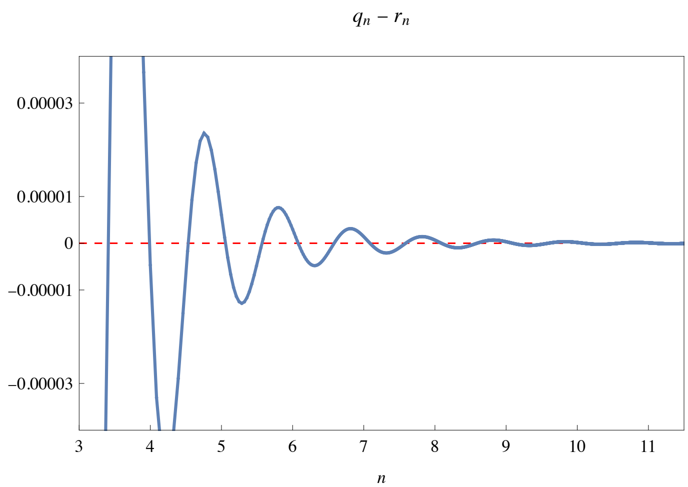
$$= \sum_{j=-k_n}^{\infty} \lambda_n\, 2^{-j-1} \left(\left(1 - \frac{1}{n}\left(\lambda_n\, 2^{-j}\right)\right)^{n - \frac{1}{\lambda_n 2^{-j}}}\right)^{\lambda_n 2^{-j}}.$$

The large expression on the right of the last expression has the form $\left(\left(1 - \frac{x_n}{n}\right)^n\right)^{\lambda_n \left(\frac{1}{2}\right)^j} \cdot \left(\left(1 - \frac{x_n}{n}\right)^{-1}\right)$ and this is asymptotic to $e^{-\lambda_n \left(\frac{1}{2}\right)^j}$ as $n \to \infty$ by the classic $\left(1 + \frac{z}{n}\right)^n \to e^z$ limit. And of course $-k_n \to -\infty$ as $n \to \infty$. So we have $q_n \sim r_n$, where

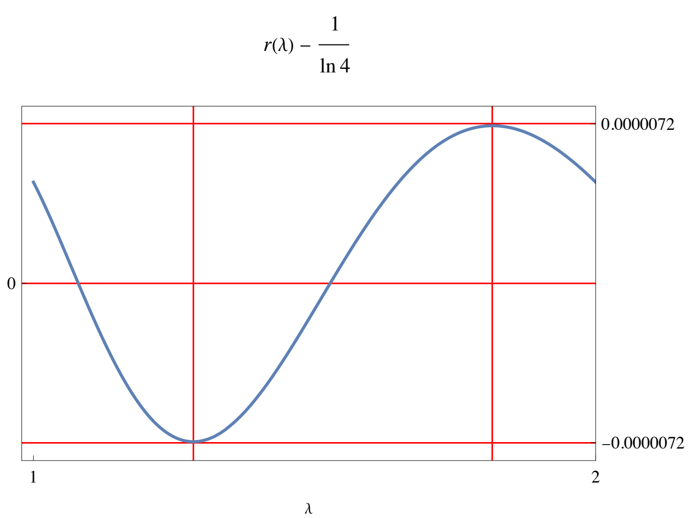
(2) $\quad r_n = \lambda_n \sum_{j=-\infty}^{\infty} \left(\frac{1}{2}\right)^{j-1} e^{-\lambda_n \left(\frac{1}{2}\right)^j}$. (One needs to justify the interchange of limits and sum.) Here is the graph of $r_n - \frac{1}{\ln 4}$ (using a log scale for the n-axis).



We can also illustrate the convergence of q_n to r_n.



The periodicity in r_n occurs because λ_n is periodic in $\log_2 n$. So to understand r_n we need examine (2) only for $1 \le \lambda \le 2$. Some easy estimation shows that letting j run from -15 to $+50$ will give a sufficiently accurate approximation to the doubly infinite sum (2). The following graph then tells the story, and periodicity implies that r_n repeats this oscillation with total amplitude of about $1.44 \cdot 10^{-5}$. The earlier graph of $r_n - \frac{1}{\ln 4}$ is just an infinite repetition of the curve below.



Because $q_n / r_n \to 1$, q_n has the same asymptotic behavior. □

1. Calkin, Canfield, and Wilf, Averaging Sequences, Deranged Mappings, and a Problem of Lambert and Slater, *J. Comb. Theory*, Ser. A **91**(1-2), 171–190 (2000);
< https://www2.math.upenn.edu/~wilf/website/LampertSlater.pdf >

2. Lennart Räde, Problem E3436, *Amer Math Monthly*, **98** (1991) 366; solution by O. P. Lossers, ibid. **101** (1994), 78.