The First Matching Pair of Socks

Andrew Ford

June 30, 2024

1 Link to problem statement

The link to the problem statement from Zach Wissner-Gross can be found at
https://thefiddler.substack.com/p/can-you-find-a-matching-pair-of-socks.

2 Original problem: 5 pairs

We wish to construct a function $f(x)$ such that $f(x)$ is the probability that the first pair of socks is found after drawing exactly x socks. In other words, the first $x-1$ socks drawn are all different, and the next sock drawn after those matches one of the previously drawn socks.

The number of ways to draw $x-1$ socks that all come from different pairs is $\binom{5}{x-1} \cdot 2^{x-1}$; the first term represents the number of pairs chosen and the second term accounts for having two available socks in each pair. There are a total of $\binom{10}{x-1}$ ways to draw $x-1$ socks without regard to pairs, so the probability that the first $x-1$ socks all come from different pairs is

$$\frac{\binom{5}{x-1} \cdot 2^{x-1}}{\binom{10}{x-1}}. \tag{1}$$

After that, there are $10-(x-1) = 11-x$ socks remaining, and $x-1$ of them match a previously drawn sock. Combining this with (1), we get

$$f(x) = \frac{\binom{5}{x-1} \cdot 2^{x-1}}{\binom{10}{x-1}} \cdot \frac{x-1}{11-x}. \tag{2}$$

This simplifies to

$$f(x) = \frac{2^{x-1}(x-1)(10-x)(9-x)(8-x)(7-x)}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}. \tag{3}$$

Obviously, we must draw at least 2 socks to get a pair. Additionally, by the pigeonhole principle, there is guaranteed to be a pair if we draw 6 socks. This means that we only need to compute $f(x)$ for $x \in \{2, 3, 4, 5, 6\}$. After plugging these values into (3), we get

x	2	3	4	5	6
$f(x)$	7/63	14/63	18/63	16/63	8/63

Therefore, the probability that the first pair will be found after drawing exactly x socks is maximized when $x = 4$.

3 Extra credit: N pairs

To start, we construct a function $g(N, x)$ to represent the probability that, with N pairs of socks in the drawer, we draw the first pair after exactly x socks. Using the same method as in the $N = 5$ case represented by

(2), we find that

$$g(N,x) = \frac{\binom{N}{x-1} \cdot 2^{x-1}}{\binom{2N}{x-1}} \cdot \frac{x-1}{2N+1-x}. \tag{4}$$

After some simplification, we can choose to rewrite this as

$$\frac{N!}{(2N)!} \cdot \frac{(2N-x)! \cdot 2^{x-1} \cdot (x-1)}{(N+1-x)!}, \tag{5}$$

separating out the terms that do not depend on x. From here, we define

$$\begin{align} h(N,x) &= \frac{g(N,x)}{g(N,x-1)} \tag{6} \\ &= \frac{\frac{N!}{(2N)!} \cdot \frac{(2N-x)! \cdot 2^{x-1} \cdot (x-1)}{(N+1-x)!}}{\frac{N!}{(2N)!} \cdot \frac{(2N-x+1)! \cdot 2^{x-2} \cdot (x-2)}{(N+2-x)!}} \tag{7} \\ &= \frac{2 \cdot (x-1) \cdot (n+2-x)}{(2N-x+1) \cdot (x-2)}. \tag{8} \end{align}$$

When $h(N,x)$ is defined in this way, $g(N,x) > g(N,x-1)$ when $h(N,x) > 1$ and $g(N,x) < g(N,x-1)$ when $h(N,x) < 1$. Since $h(N,x)$ is continuous with respect to x when $x \in (1, 2N-1)$, and it is only reasonable for the probability to be nonzero when $x \in \{2, 3, \ldots, N+1\}$, it is sufficient to solve $h(N,x) = 1$ for x in terms of N to determine candidates for the value of x which maximizes $g(N,x)$.

By (8), we have

$$\begin{align} 1 &= \frac{2 \cdot (x-1) \cdot (N+2-x)}{(2N-x+1) \cdot (x-2)} \tag{9} \\ (2N-x+1) \cdot (x-2) &= 2 \cdot (x-1) \cdot (N+2-x) \tag{10} \\ 2Nx - x^2 + x - 4N + 2x - 2 &= 2Nx + 4x - 2x^2 - 2N - 4 + 2x \tag{11} \\ x^2 - 3x - 2N + 2 &= 0 \tag{12} \\ x &= \frac{3 \pm \sqrt{8N+1}}{2}. \tag{13} \end{align}$$

Since N is assumed to be "very large" in the problem, the solution $\frac{3-\sqrt{8N+1}}{2}$ yields a negative value of x. Since only the positive solution makes sense within the constraints of the problem, it follows that

$$h\left(N, \left\lceil \frac{3+\sqrt{8N+1}}{2} \right\rceil\right) < 1 < h\left(N, \left\lfloor \frac{3+\sqrt{8N+1}}{2} \right\rfloor\right), \tag{14}$$

which means the probability that the first pair is drawn after exactly x socks are pulled from the drawer is maximized when

$$x = \left\lfloor \frac{3+\sqrt{8N+1}}{2} \right\rfloor. \tag{15}$$