Pairs of socks

Arjen Markus

September 27, 2023

Introduction

Consider the following simple combinatorial problem: you have a drawer full of socks and you blindly draw a number of them. What is the probability of drawing a complete pair? What happens to the probabibilities if you need three of a kind or four of a kind?

We assume that the pairs of socks are all complete, similarly for the objects that come in three or four or ... Of course, we might also consider the case where there is a only a single sock of a pair, but that is beyond the current investigation.

Straightforward observations

If you have n sets of items, then drawing $n + 1$ items guarantees that you have at least one complete pair (or in general: one complete set). But the larger the number of sets. the more unlikely it will become to get a complete set with a smaller number.

Let us consider two pairs of socks for the moment, a pair of red socks and a pair of blue socks. Then taking two at random from the drawer leads to three possibilities:

- Two red socks
- Two blue socks
- One red and one blue sock. Since there are two such mixed pairs, this possibility is twice as likely.

Hence, we can conclude that in this case drawing a complete pair has a probability of $\frac{1}{2}$. Another line of reasoning, however, leads to $\frac{1}{3}$:

- Draw the first sock, it is either red or blue, that does not matter.
- We are left with one sock of the right colour and two socks of the wrong colour.

- The probability of drawing a sock of the right colour, and thereby obtaining a complete pair, is therefore $\frac{1}{3}$.

I have no explanation for this discrepancy, other than that combinatorial problems can surprise us. In any case, with the general cases, it is easier and more reliable to use combinatorial formulae and visualisation techniques.

The general case: n pairs of socks

If we have n pairs, then the total number of socks is $2n$, so that the number N of possible combinations becomes:

$$\begin{aligned} N &= \binom{2n}{n} \\ &= n(2n-1) \end{aligned}$$

The number of combinations that have a complete pair is, of course, n, as there are n pairs. So, the probability of getting a complete pair is $1/(2n-1)$.

Now, what happens if we draw m socks in stead of 2?

We again have the total number of combinations and the number of acceptable combinations – even if we draw two socks of the same pair rightaway, we can continue drawing socks up to the total of m. The rest does not matter, but the counting is much simpler.

The total number of combinations is:

$$N = \binom{2n}{m}$$

Naïve reasoning leads to us to conclude that the number of acceptable combinations M is:

$$M = n\binom{2n-2}{m-2}$$

as the $m-2$ extra socks that are drawn from the remaining $2n-2$ do not matter.

To illustrate this, consider the case of three pairs of socks, numbered 1 to 6. Socks 1 and 2 form a pair, 3 and 4 are the second pair and 5 and 6 the third. Then the combinations possible with three arbitrary socks are:

1, 2, 3 – pair 1
1, 2, 4 – pair 1
1, 2, 5 – pair 1
1, 2, 6 – pair 1
1, 3, 4 – pair 2
1, 3, 5
1, 3, 6
1, 4, 5
1, 4, 6

1, 5, 6 – pair 3
2, 3, 4 – pair 2
2, 3, 5
2, 3, 6
2, 4, 5
2, 4, 6
2, 5, 6 – pair 3
3, 4, 5 – pair 2
3, 4, 6 – pair 2
3, 5, 6 – pair 3
4, 5, 6 – pair 3
Of the total of 20 combinations 12 contain a pair.
In general, the probability p of retrieving at least one pair is:

$$p \quad = \quad n\binom{2n-2}{m-2} \Big/ \binom{2n}{m}$$

Filling in $n = 3, m = 3$, we get:

$$p \quad = \quad 3\binom{4}{1} \Big/ \binom{6}{3} = 3 \cdot 4/20 = \frac{3}{5}$$

Filling in $n = 2, m = 3$, we get:

$$p \quad = \quad 2\binom{2}{1} \Big/ \binom{4}{3} = 2 \cdot 2/4 = 1$$

But filling in $n = 2, m = 4$, we get:

$$p \quad = \quad 2\binom{2}{2} \Big/ \binom{4}{4} = 2 \cdot 1/1 = 2 > 1$$

What we have overlooked is the fact that with more than three socks to draw, sometimes we will have a combination with two or even more complete pairs. Hence we counting several combinations twice or three times.

To solve this problem we need a different approach: let us look at the combinations that certain do not contain a single complet pair and subtract that from the total number of combinations.

Observe that drawing the first sock will give us an incomplete pair and there are $2n$ possibilities to do so. Drawing the second sock should be done on the set where the matching sock is not included. Thus, we have $2n - 2$ possibilities for drawing the next sock. For the third sock we need again to exclude the remaining matching sock, so we get to draw from $2n - 4$ socks, and so on. With this process the number of possibilities relies on the permutation of all these single socks, so we need to correct for the number of permutations of m objects.

In general the number M_* of ways to draw m socks without forming a pair is:

$$
\begin{aligned}
M_* &= 2n \cdot (2n-2) \cdot (2n-4) \ldots (2n-2m+2)/m! \\
&= 2^m \binom{n}{m} \qquad (1)
\end{aligned}
$$

(A quick check on the correctness of this formula: selecting one sock from $2n$ socks gives $2n$ possibilities – $2 \cdot \binom{n}{1} = 2n$)

Hence the formula for the probability of getting at least one pair by drawing m socks from n pairs is:

$$
p = 1 - 2^m \binom{n}{m} \Big/ \binom{2n}{m}
$$

How many socks to draw?

With the formula we can estimate how many socks you would have to draw to be reasonably sure that there is at least one complete pair among them. Let us define this as a probability equal to or larger than $\frac{1}{2}$.

The table below gives the estimate found in this way.

Table 1: The number of socks to draw for a probability $\geq \frac{1}{2}$.

n	m	n	m
2	3	9	5
3	3	10	6
4	4	15	7
5	4	20	8
6	4	25	9
7	5	30	9
8	5	100	17

Extension to sets of three or more objects

If, instead of pairs, we have objects that come in triples or quadruples, the formulae will change, but the line of reasoning will not.

So, the number of combinations N when selecting m objects from $3n$ triples, is (m must be at least 3, as otherwise we cannot have a full triplet):

$$
N = \binom{3n}{m}
$$

The number of combinations that are guaranteed not to contain a full triplet is:

$$\begin{aligned} M_* &= 3n \cdot (3n-3) \cdot (3n-6)...(3n-3m+3)/m! \\ &= 3^m \binom{n}{m} \end{aligned} \tag{2}$$

And therefore the probability of drawing a full triplet from a set of n triplets at random is:

$$p = 1 - 3^m \binom{n}{m} \Big/ \binom{3n}{m}$$

The general case of k-tuplets:

$$N = \binom{kn}{m}$$

$$\begin{aligned} M_* &= kn \cdot (kn-k) \cdot (kn-2k)...(3n-km+k)/m! \\ &= k^m \binom{n}{m} \end{aligned} \tag{3}$$

And therefore the probability of drawing a full triplet from a set of n triplets at random is:

$$p = 1 - k^m \binom{n}{m} \Big/ \binom{kn}{m}$$