The Workshop "*ICEBREAKER EVENT*" problem

Des McLernon[1]*, Ali Hayajneh*[2]*, and Ali Zaidi*[3]

[1]*School of Electronic and Electrical Engineering, The University of Leeds, Leeds, LS29JT, United Kingdom,* d.c.mclernon@leeds.ac.uk

[2]*Department of Electrical Engineering, Faculty of Engineering, Hashemite University, Zarqa 13133, Jordan,* alihayajneh@hu.edu.jo

[3]*School of Electronic and Electrical Engineering, The University of Leeds, Leeds, LS29JT, United Kingdom ,* s.a.zaidi@leeds.ac.uk

Abstract

We assume that N different people have N different name badges. Let the N name badges be randomly redistributed among the N individuals so that nobody now has their own (correct) name badge (i.e., a *derangement* of N objects). We define $P(N)$ as the probability that after the derangement at least one pair of individuals will have swapped their name badges (i.e., one or more "*swapped pairs*"). We will show that:

a) $P(N) = -\dfrac{N!}{!N}\displaystyle\sum_{m=1}^{\lfloor N/2\rfloor}\left(-\frac{1}{2}\right)^m \frac{1}{m!}\frac{!(N-2m)}{(N-2m)!},\ N \geqslant 2, \quad \lim_{N\to\infty} P(N) = 1 - \frac{1}{\sqrt{e}} \approx 0.393469$ where $!M = M!\displaystyle\sum_{i=0}^{M}\frac{(-1)^i}{i!}$ represents the number of possible derangements of M objects.

b) In addition, if we define $K(N) \geqslant 0$ as the discrete random variable that gives the number of swapped pairs in any randomly chosen derangement of N objects, then $\text{Prob}(K(N) = k) \approx \dfrac{0.5^k e^{-0.5}}{k!}, \quad k = 0, 1, 2, \ldots,\ \overline{K(N)} = 0.5\left(\dfrac{N!}{!N}\right)\left(\dfrac{!(N-2)}{(N-2)!}\right)$ and $\lim_{N\to\infty}\overline{K(N)} = E(K(N)) = 0.5$.

1 Introduction

1.1 Motivation

The original motivation for examining this problem was as follows:

©2026 Des McLernon, Ali Hayajneh and Ali Zaidi. This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

> "We recently attended a technical workshop, and the first session was '*the icebreaker event*'. Every one of the N attendees from around the world had a name badge that was put into a hat and the organiser then (randomly) returned a badge to each attendee while ensuring that nobody got their own badge back. You then had to search out the person whose badge you were given and report back to the workshop about your conversation with that person. We noticed that there were two pairs of attendees where the attendees in each pair had exchanged badges with each other (i.e., two pair-swaps). So we wondered "*What is the formula for the probability (P(N)) that at least one pair of the N attendees will have exchanged badges with each other*?"

But you might also ask, why does it even matter? So, consider for example, that just a single pair-swap has occurred.

Well the format for the icebreaker event was that after all the badges were exchanged and each attendee had searched out the owner of the badge that they had now been given, a random attendee's name (say John) was called out by the organiser. John then reported that he had, say, Margaret's badge. He then described all that he had learnt about Margaret in his conversation with her. Then Margaret took to the floor and said that she had been given Abdul's badge. She then described all that she had learnt about Abdul. Abdul then took to the floor and said that he had Pablo's badge and so on, until all the N attendees had their say and everyone had been effectively introduced to everyone else in the room.

But what if, say, attendee Alice was given Manuel's badge and Manuel was given Alice's badge (i.e., a pair-swap)? Then the icebreaker event would either (i) finish with just these two attendees reporting back if either Alice or Manuel was called first, or (ii) fail to include either of them if someone else was called first.

1.2 A Mathematical Reformulation into an Equivalent Generic Problem

We start with N objects $(B_j,\ j = 1, 2, \ldots, N)$ in N labelled containers and where object B_j was originally in container 'j'. We redistribute the objects (with one placed in each container) such that no object returns to its original container. This is called a '*derangement*'. We are interested in the formula for the probability $P(N)$ that one pair of objects (or more than one pair) will have effectively swapped containers. For example, if in one derangement object B_3 ends up in container 6 and object B_6 ends up in container 3, then we say that '*one pair-swap has occurred*'. Alternatively, if in another derangement object B_2 ends up in container 4 and object B_4 ends up in container 2, while object B_{21} ends up in container 14 and object B_{14} ends up in container 21, then we say that '*two pair-swaps have occurred*', etc. A '*pair-swap*' in a derangement is also more formally called a '*2-cycle*'.

1.3 Nomenclature

Note that a derangement of N objects is a classical problem where N objects are re-ordered so that no object appears in its original position. The number of possible derangements is usually written as $!N$ where [Wiki]:

$$!N = N! \sum_{i=0}^{N} \frac{(-1)^i}{i!}, \ N = 0, 1, 2, 3, \ldots = \left[\frac{N!}{e} + 0.5 \right], \ N = 1, 2, 3, \ldots \tag{1}$$

Now let $D_N(m), \ m = 0, 1, 2, \ldots$ represent the number of derangements of $N \geqslant 2$ objects (from a total of $!N$ possible derangements) where *exactly m pairs of objects have swapped containers*. For example, consider $N = 4$ objects below.

Container 1	**Container 2**	**Container 3**	**Container 4**
B_2	B_1	B_4	B_3
B_2	B_3	B_4	B_1
B_2	B_4	B_1	B_3
B_3	B_1	B_4	B_2
B_3	B_4	B_1	B_2
B_3	B_4	B_2	B_1
B_4	B_1	B_2	B_3
B_4	B_3	B_1	B_2
B_4	B_3	B_2	B_1

Table 1: All nine possible derangements of $N = 4$ objects with 'pair-swaps' colour-coded.

Note that in the first derangement objects 1 and 2 swap containers (green) and objects 3 and 4 (red) do likewise. Here we say two pairs have swapped. Similarly for derangements five and nine where two pairs also swap. The remaining six derangements have no pair-swaps. Thus we get $D_4(0) = 6$ and $D_4(2) = 3$. Note also that we cannot have a one pair-swap because then the remaining two objects will be in their original positions and so we will not have a derangement. So for $N = 4$, the allowable values for 'm' to get non-zero $D_4(m)$ are just 0 and 2, with all other (non-negative) integer values of 'm' giving $D_4(m) = 0, \ m = 1, 3, 4, 5, \ldots$. With a little thought we can see that the allowable values for 'm' to get non-zero $D_N(m)$ are as follows, using $2 \leqslant N \leqslant 9$ as an illustration:

$$\begin{aligned}
N = 2 &: m = 1 \\
N = 3 &: m = 0 \\
N = 4 &: m = 0, 2 \\
N = 5 &: m = 0, 1 \\
N = 6 &: m = 0, 1, 3 \\
N = 7 &: m = 0, 1, 2 \\
N = 8 &: m = 0, 1, 2, 4 \\
N = 9 &: m = 0, 1, 2, 3
\end{aligned} \tag{2}$$

Thus we can write the range for 'm' that gives non-zero $D_N(m)$ below (with all other 'm' giving $D_N(m) = 0$):

$$\begin{aligned}
N = 2: &\quad m = 1 \\
\text{Odd } N \geqslant 3: &\quad m = 0, 1, 2, \ldots, \frac{N-3}{2} \\
\text{Even } N \geqslant 4: &\quad m = 0, 1, 2, \ldots, \left(\frac{N}{2} - 3\right), \left(\frac{N}{2} - 2\right), \frac{N}{2}
\end{aligned} \tag{3}$$

So while we could write $D_N(m),\ m = 0, 1, \ldots, \infty$ to pick up every zero and non-zero valued $D_N(m)$ we will just write for a finite range of 'm' and $D_N(m),\ m = 0, 1, \ldots, \lfloor N/2 \rfloor$, which will pick up at most just one $D_N(m) = 0$.

Now let us define $D_N = \sum_{m=1}^{\lfloor N/2 \rfloor} D_N(m)$ as the total number of derangements of N elements (from a total of $!N$ possible derangements) where there are one or more pair exchanges/pair-swaps in each derangement. Thus:

$$P(N) = \frac{D_N}{!N} = \frac{\sum_{m=1}^{\lfloor N/2 \rfloor} D_N(m)}{N! \sum_{i=0}^{N} (-1)^i / i!}, \quad N = 2, 3, 4, \ldots \tag{4}$$

Also, define $K(N)$ as the random variable that gives the number of swapped pairs in any randomly chosen derangement of N objects where its mean value is $\overline{K(N)} = E(K(N))$.

So as an example we then have from Table 1, that where $N = 4$, then $P(4) = \frac{D_4}{!4} = \frac{D_4(1)+D_4(2)}{!4} = \frac{0+3}{9} = \frac{1}{3}$ and $\overline{K(4)} = \left(\frac{D_4(0)}{!4}\right) 0 + \left(\frac{D_4(1)}{!4}\right) 1 + \left(\frac{D_4(2)}{!4}\right) 2 = \frac{2}{3}$.

1.4 Objectives

So we want to answer the following questions for any $N \geqslant 2$:

(a) What are the closed-form expressions for $P(N)$ and $\overline{K(N)}$?

(b) What are $\lim_{N\to\infty} P(N)$ and $\lim_{N\to\infty} \overline{K(N)}$?

(c) What is the probability mass function (pmf) for $K(N)$?

2 A recursive formula for $D_N(m)$

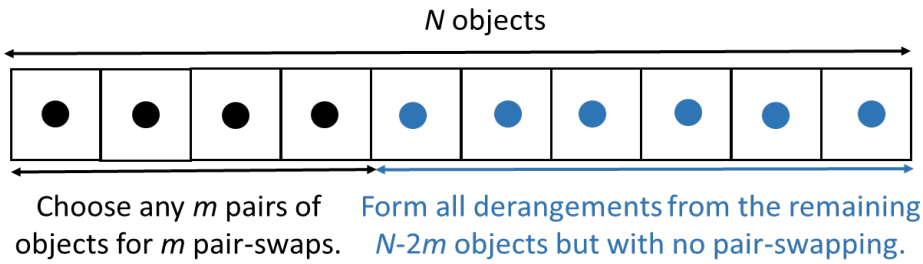


Figure 1: Constructing all the m pair-swap derangements of N objects ($D_N(m)$ in total).

From the definition for $D_N(m)$ we want to construct all the derangements of N objects in N containers where exactly m pairs of objects swap containers – that is, each of the two elements in any pair will swap positions, see Fig. 1. So we start by choosing m object pairs from N objects (see black circles in Fig. 1). It is not difficult to show that this can be done in $a(m) = \frac{{}^{N}C_2\,{}^{N-2}C_2 \ldots {}^{N-2(m-1)}C_2}{m!} = \frac{N!}{2^m m!(N-2m)!}$ ways with $m = 1, 2, 3, \ldots$ and where ${}^{n}C_r = \frac{n!}{(n-r)!r!}, \quad r = 0, 1, \ldots, n.$

The remaining N-2m objects (see blue circles in Fig. 1) must then form $!(N\text{-}2m)$ derangements to construct all the allowable derangements of N objects with m pair-swaps. But some of these $!(N\text{-}2m)$ derangements may also include one or more pair-swaps. This cannot be allowed as we are trying to construct all the N-element derangements where exactly m pairs swap, and we have already chosen the m (black) pairs that are to swap (see black circles in Fig. 1).

Thus we need to now only consider what we will call the '*pure derangements*' of these remaining N-2m blue objects in Fig. 1 – i.e., they must be restricted to zero pair-swaps. So we have to subtract from all these possible $!(N\text{-}2m)$ derangements, all of the derangements that have one pair-swap or two pair-swaps, or three pair-swaps, etc. From the definition for $D_N(m)$ that gives the term $!(N-2m) - \sum_{k=1}^{\lfloor (N-2m)/2 \rfloor} D_{N-2m}(k)$ as the total number of zero pair-swap derangements from N-2m objects. Thus we can write:

$$D_N(m) = a(m)\left(!(N-2m) - \sum_{k=1}^{\lfloor (N-2m)/2 \rfloor} D_{N-2m}(k)\right), \quad a(m) = \frac{N!}{2^m m!(N-2m)!} \tag{5}$$

We can then insert (5) into (4) to get $P(N)$. But the big problem here is that this is not a closed-form expression and so $D_N(m)$ may need to be calculated recursively. Here are two examples.

Example 2.1.

$$D_9(3) = a_9(3)\left(!3 - \sum_{k=1}^{1} D_3(k)\right) = 1260(2-0) = 2520$$

where $D_3(1) = 0$.

Example 2.2. *(where only non-zero* $D_M(k)$ *are shown in the calculations below):*

$$D_9(1) = a_9(1) \times \left(!(7) - \sum_{k=1}^{3} D_7(k)\right) = a_9(1) \times (!(7) - D_7(1) - D_7(2)) = 41,040$$

where

$$D_7(1) = a_7(1) \times \left(!(5) - \sum_{k=1}^{2} D_5(k)\right) = a_7(1) \times (!(5) - D_5(1)) = 504,\ D_5(2){=}0$$

$$D_5(1) = a_5(1) \times \left(!(3) - \sum_{k=1}^{1} D_3(k)\right) = a_5(1) \times (!(3) - 0) = a_5(1){\times}!(3) = 20,\ D_3(1){=}0$$

$$D_7(2) = a_7(2) \times \left(!(3) - \sum_{k=1}^{1} D_3(k)\right) = a_7(2) \times (!(3) - 0) = a_7(2){\times}!(3) = 210,\ D_3(1){=}0$$

So while (4) *and* (5) *can indeed generate values for* P(N)*, because* (5) *is recursive it is not very helpful in calculating closed-form expressions for* $\lim_{N\to\infty} P(N)$, $\lim_{N\to\infty} \overline{K(N)}$ *or the pmf for* $K(N)$.

3 A non-recursive formula for $D_N(m)$ and $P(N)$

Now we can easily make (5) non-recursive if we define a new variable $Z_N(m)$ where we drop the (non pair-swapping) constraint on the derangements of the $N - 2m$ objects in (5). It is intuitively appealing to do that, and so we get:

$$\underbrace{Z_N(m)}_{\substack{\text{number of ways of choosing } m \text{ pairs from } N \text{ objects that will swap the while remaining } (N{-}2m) \text{ objects form a (non-pure) derangement with no restriction on whether it includes other swapped pairs.}}} = \underbrace{\left(\frac{N!}{2^m m!(N-2m)!}\right)}_{\substack{\text{number of ways to choose } m \text{ pairs from } N \text{ objects.}}} \times \underbrace{!(N-2m)}_{\substack{\text{number of unconstrained derangements for } N{-}2m \text{ objects that will include one or more pair swaps.}}}, \ m = 1,2,3,\ldots$$

$$\implies Z_N(m) = \left(\frac{N!}{2^m m!}\right)\left(\frac{!(N-2m)}{(N-2m)!}\right), \quad m = 1,2,3,\ldots \tag{6}$$

But how do we now derive $D_N(m)$ from the non-recursive expression for $Z_N(m)$ in (6), and is it even possible? Without loss of generality, and to ease nomenclature, let us look at the particular case of $N = 13$ for $Z_{13}(m)$. But we will first start with $Z_{13}(2)$ as in Fig. 2 to establish the method.

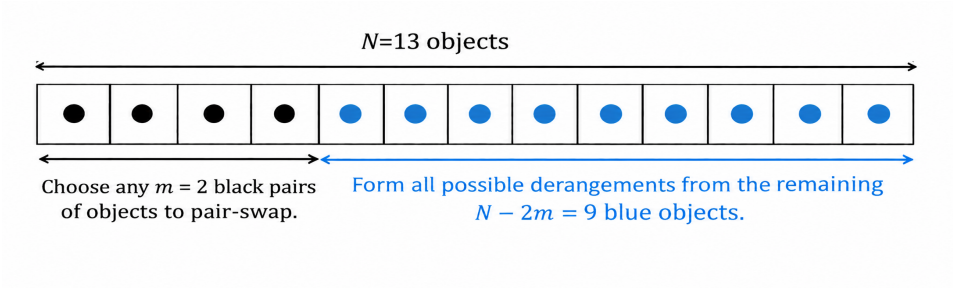


Figure 2: Looking at all the possible N=13 object derangements contributing to $Z_{13}(2)$ in (6).

We first choose two pairs of objects (in black) to swap positions. The remaining 9 blue objects are then deranged without any constraint (as in (6) to form !9 derangements. These !9 derangements will obviously include 9 object derangements with zero pair-swaps, one pair-swap, two pair-swaps and three pair-swaps. Combined with the chosen two black pair-swaps, this will produce overall, 13 object derangements with different numbers of two pair-swaps, three pair-swaps, four pair-swaps and five pair-swaps. Let us now look at the contribution from each case in turn.

From the definitions of $D_N(m)$ and $Z_N(m)$, then we can start by writing

$$Z_{13}(2) = \sum_{j=1}^{5} A_{13}(2, j) D_{13}(j) \tag{7}$$

where $\{A_{13}(2, j)\}_{j=1}^{5}$ are the unknown coefficients that we need to find and $D_{13}(j) \neq 0$ for this range.

(i) Consider first, $A_{13}(2, 1)$ - the contribution of 13 object derangements with single pair-swaps included in $Z_{13}(2)$. Now as we always have a minimum of two pairs swapping (see black objects in Fig. 2), and the minimum m pair-swap for the 9 blue objects is a zero pair-swap, then there can be no overall 13 object derangements with single pair-swaps (i.e., $D_{13}(1)$) included in $Z_{13}(2)$. So $A_{13}(2, 1) = 0$ in (7).

(ii) Consider now $A_{13}(2, 2)$ - the contribution of 13 object derangements with double pair-swaps included in $Z_{13}(2)$. Now the !9 blue object derangements will include a subset with zero pair-swaps. These will combine with the (black) double pair-swap in Fig. 2, to give $D_{13}(2)$ 13 object double pair-swaps. So clearly $A_{13}(2, 2) = 1$ in (7).

(iii) Consider now $A_{13}(2, 3)$ - the contribution of 13 object derangements with three pair-swaps included in $Z_{13}(2)$. Now the !9 blue object derangements in Fig. 2 will

include a subset with single pair-swaps. These will combine with the (black) double pair-swap in Fig. 2, to give a number of 13 object three pair-swaps. So let objects 1 and 2, 3 and 4, along with 12 and 13 in Fig. 2 all swap positions (i.e., three pair swaps in total). This is illustrated in Fig. 3 below.

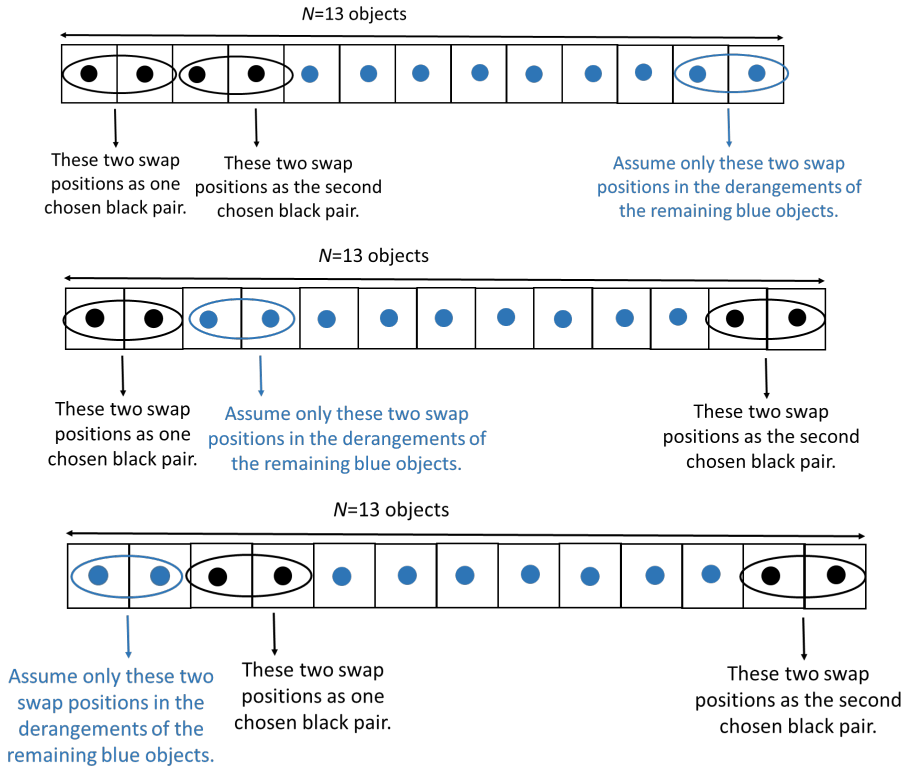


Figure 3: The three different ways that a specified triple pair-swap can contribute to $Z_{13}(2)$.

So the problem now effectively becomes – assume we have three (uncoloured) named pairs of objects (1 and 2, 3 and 4, 12 and 13) and we want to colour two pairs black and the remaining pair blue, thus leading to three pair swaps. This can be done in ${}^3C_2$ ways thus contributing ${}^3C_2\, D_{13}(3)$ 13 object derangements with three pair-swaps to $Z_{13}(2)$. So $A_{13}(2,3) = {}^3C_2$ in (7).

(iv) With similar reasoning to above, it is not difficult to show that $A_{13}(2,4) = {}^4C_2$ and $A_{13}(2,5) = {}^5C_2$ and thus (7) becomes

$$Z_{13}(2) = \sum_{j=2}^{5} {}^jC_2 D_{13}(j) \tag{8}$$

(v) Repeating this analysis for all other values of m in $Z_{13}(m)$ and we get the final

result that

$$Z_{13}(m) = \sum_{j=m}^{5} {}^{j}C_m D_{13}(j), \quad m = 1, 2, 3, 4, 5$$

which can be re-written as

$$\mathbf{z}_{13} = \begin{bmatrix} Z_{13}(1) \\ Z_{13}(2) \\ Z_{13}(3) \\ Z_{13}(4) \\ Z_{13}(5) \end{bmatrix} = \begin{bmatrix} {}^{1}C_1 & {}^{2}C_1 & {}^{3}C_1 & {}^{4}C_1 & {}^{5}C_1 \\ 0 & {}^{2}C_2 & {}^{3}C_2 & {}^{4}C_2 & {}^{5}C_2 \\ 0 & 0 & {}^{3}C_3 & {}^{4}C_3 & {}^{5}C_3 \\ 0 & 0 & 0 & {}^{4}C_4 & {}^{5}C_4 \\ 0 & 0 & 0 & 0 & {}^{5}C_5 \end{bmatrix} \begin{bmatrix} D_{13}(1) \\ D_{13}(2) \\ D_{13}(3) \\ D_{13}(4) \\ D_{13}(5) \end{bmatrix} = \mathbf{A}_{13}\mathbf{d}_{13} \tag{9}$$

And if we require $\{D_{13}(m)\}_{m=0}^{5}$, then from (9), and the definition of $D_{13}(m)$, we get

$$\mathbf{d}_{13} = \mathbf{A}_{13}^{-1}\mathbf{z}_{13}, \quad D_{13}(0) = !N - \sum_{m=1}^{5} D_{13}(m), \quad Z_N(m) = \left(\frac{N!}{2^m m!}\right)\left(\frac{!(N-2m)}{(N-2m)!}\right) \tag{10}$$

This non-recursive evaluation of $D_{13}(m)$ will give the same result as the recursive method in (5). But we also now have

$$D_{13} = \sum_{m=1}^{5} D_{13}(m) = [1\ 1\ 1\ 1\ 1]\mathbf{d}_{13} = [1\ 1\ 1\ 1\ 1]\mathbf{A}_{13}^{-1}\mathbf{z}_{13} \tag{11}$$

for use in (4) to get $P(13)$. But while non-recursive, it is still not a simple expression to play around with. However from the easily proven result that $\sum_{i=1}^{k}(-1)^{i+1k}C_i = 1, \quad k = 1, 2, 3, \ldots$ (see Appendix A1), along with the definition of $\mathbf{A}_{13}$ in (9), then we can show that

$$[1\ 1\ 1\ 1\ 1] = [1\ -1\ 1\ -1\ 1]\mathbf{A}_{13} \Rightarrow [1\ 1\ 1\ 1\ 1]\mathbf{A}_{13}^{-1} = [1\ -1\ 1\ -1\ 1] \tag{12}$$

Now substituting (12) into (11) gives

$$D_{13} = \sum_{m=1}^{5} D_{13}(m) = [1\ -1\ 1\ -1\ 1]\mathbf{z}_{13} = \sum_{m=1}^{5}(-1)^{m+1}Z_{13}(m) \tag{13}$$

So in general (from (13) and (10) we can easily show that for any N:

$$D_N = \sum_{m=1}^{\lfloor N/2 \rfloor} D_N(m) = \sum_{m=1}^{\lfloor N/2 \rfloor}(-1)^{m+1}Z_N(m), \quad Z_N(m) = \left(\frac{N!}{2^m m!}\right)\left(\frac{!(N-2m)}{(N-2m)!}\right) \tag{14}$$

(Please note that putting N=13 in (14) and getting $\sum_{m=1}^{6}$ is still consistent with $\sum_{m=1}^{5}$ in (13) as $Z_{13}(6) = 0$ in (14).)

So finally, a closed-form expression for the probability $P(N)$ "*that one or more pairs of objects will have effectively swapped containers after derangement*" is from (4) and (14):

$$P(N) = \frac{D_N}{!N} = -\sum_{m=1}^{\lfloor N/2 \rfloor} \left(-\frac{1}{2}\right)^m \frac{1}{m!} \left(\frac{N!}{!N}\right) \left(\frac{!(N-2m)}{(N-2m)!}\right), \quad N \geqslant 2 \tag{15}$$

Note that $P(N)$ agrees with computer simulations. Now from [Wiki] $\lim_{N\to\infty} \frac{N!}{!N} = e$ and so $\lim_{N\to\infty} \frac{!(N-2m)}{(N-2m)!} = \frac{1}{e}$. And as the limit of a product is the product of the limits (if limits are finite and exist) and since $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, then from (15) we can say that

$$\lim_{N\to\infty} P(N) = -\sum_{m=1}^{\infty} \left(-\frac{1}{2}\right)^m \frac{1}{m!} = 1 - \frac{1}{\sqrt{e}} \approx 0.393469 \tag{16}$$

Note that this limit in (16) is approached very quickly (see (17) below). So that explains why, at the workshop icebreaker event previously described (approx. N = 40 attendees), my having observed one or more name badge pair-swaps is not a very unlikely event.

$$\begin{aligned} P(2) &= 1 \\ P(3) &= 0 \\ P(4) &= 0.3333 \\ P(5) &= 0.4545 \\ P(6) &= 0.3962 \\ P(7) &= 0.3851 \\ P(8) &= 0.3941 \\ P(9) &= 0.3943 \end{aligned} \tag{17}$$

4 Expressions for $\overline{K(N)} = E(K(N))$ and pmf $\mathrm{Prob}(K(N) = k)$

There is an additional interesting result for the random variable $K(N)$, the number of swapped pairs in any derangement of N elements. For example, if $N = 9$ then $K(9)$ is a discrete random variable that can only have four values 0,1,2 and 3. But what is its expected value $\overline{K(N)} = E(K(N))$ for any N? That is, what is the average number of pair-swaps for any N object derangement?

The solution is as follows. It is clear that for N objects there are in total $!N$ possible derangements. From those $!N$ possible derangements, the number of derangements with

exactly k pair-swaps in each derangement is $D_N(k),\ k \geqslant 0$. So the probability of having k pair swaps in a random derangement is:

$$\text{Prob}(K(N) = k) = \frac{D_N(k)}{!N}, \quad k = 0, 1, 2, \ldots, \lfloor N/2 \rfloor$$

$$\implies \overline{K(N)} = E(K(N)) = \sum_{k=0}^{\lfloor N/2 \rfloor} k \left(\frac{D_N(k)}{!N} \right) = \frac{1}{!N} \sum_{k=1}^{\lfloor N/2 \rfloor} k D_N(k) \tag{18}$$

Now we cannot use $D_N(k)$ from (5) in (18) because of its recursive nature. Likewise, the equivalent $D_N(k)$ from (10) is also too complicated. But fortuitously we do not actually require $D_N(k)$ explicitly in (18), but only $\sum_{k=1}^{\lfloor N/2 \rfloor} k D_N(k)$. Now from (9) we can see that $Z_N(1) = \sum_{k=1}^{\lfloor N/2 \rfloor} k D_N(k)$ and so from (6) and (18) we get the exact result:

$$\overline{K(N)} = \frac{Z_N(1)}{!N} = 0.5 \left(\frac{N!}{!N} \right) \left(\frac{!(N-2)}{(N-2)!} \right) \tag{19}$$

Using the same argument for limits as was done in (15) then from (19) we have the final result:

$$\lim_{N \to \infty} \overline{K(N)} = 0.5 \tag{20}$$

And if (19) then is evaluated for different values of N we can see in (21) below that $\overline{K(N)}$ approaches 0.5 very quickly for increasing N:

$$\begin{aligned}
\overline{K(2)} &= 1 \\
\overline{K(3)} &= 0 \\
\overline{K(4)} &= 0.667 \\
\overline{K(5)} &= 0.455 \\
\overline{K(6)} &= 0.509 \\
\overline{K(7)} &= 0.493 \\
\overline{K(8)} &= 0.500 \\
\overline{K(9)} &= 0.500
\end{aligned} \tag{21}$$

Finally, like many counting/combinatorial problems, it is an intuitive guess that the random variable $K(N)$ should approach a Poisson pmf for large N. Let us now test that hypothesis without a formal proof. For a general Poisson distributed rv X we can say:

$$\text{Prob}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad E(X) = \lambda, \quad k = 0, 1, 2, 3, \ldots \tag{22}$$

Now since we have $E(K(N)) \to 0.5$, as $N \to \infty$ then for large N we can suppose that (based on the Poisson pmf conjecture):

$$\text{Prob}(K(N) = k) \approx \frac{0.5^k e^{-0.5}}{k!}, \quad k = 0, 1, 2, \ldots \tag{23}$$

First note that from (23) we can also say that for large N: $P(N) \approx 1 - \text{Prob}(K(N) = 0) = 1 - \frac{1}{\sqrt{e}}$, which is pleasingly consistent with (16). Now the exact pmf for $K(N)$ is

$$\text{Prob}(K(N) = k) = \frac{D_N(k)}{!N}, \quad k = 0, 1, 2, \ldots \tag{24}$$

where $D_N(k)$ can be obtained from either simulation, (5) or (10).

So as an example, in Table 2, for the particular case of N=11, we have compared (24) (the exact pmf of $K(11)$) with (23) (the approximate Poisson pmf model for $K(11)$).

	$\text{Prob}(K(11) = k) = \frac{D_{11}(k)}{!11}$ (exact)	$\text{Prob}(K(11) = k) \approx \frac{0.5^k e^{-0.5}}{k!}$ (approx.)
$k = 0$	0.606587731203569	0.606530659712633
$k = 1$	0.302870291741604	0.303265329856317
$k = 2$	0.076856183054730	0.075816332464079
$k = 3$	0.011326174344908	0.012636055410680
$k = 4$	0.002359619655189	0.001579506926335
$k = 5$	0	0.000157950692633

Table 2: Close agreement of exact pmf for $K(11)$ (see (24)) with the approximate Poisson model (see (23)).

And so the previous conjecture appears to be true, that the number of swapped pairs ($K(N)$) in any derangement of N elements has a very good approximation to a Poisson random variable for large N. Finally, Table 4 in the Appendix that follows summarises some of the results presented in this paper.

5 Conclusions

The particular mathematical problem of looking at the derangement of N objects and the number of pair-swaps has been solved with simple closed-form expressions, and all the objectives in section 1.4 have been achieved. These include:

(a) Closed form expressions for D_N, $P(N)$ and $\overline{K(N)}$ in (14), (15) and (19) respectively;

(b) Expressions for $\lim_{N\to\infty} P(N) = 1 - \frac{1}{\sqrt{e}} \approx 0.393469$ and $\lim_{N\to\infty} \overline{K(N)} = 0.5$ in (16) and (20) respectively;

(c) A good Poisson approximation for the pmf of $K(N)$ in (23) - Prob$(K(N)=k) \approx \frac{0.5^k e^{-0.5}}{k!}, \quad k=0,1,2,\ldots$

Now the following comments are relevant and important.

(i) Note that an alternative non closed-form expression for D_N (and for $P(N)$ in (4)) can also be obtained by using *Exponential Generating Functions* (EGFs), but it is still more complicated than (14) and (15) and so I have not fully derived the theory here.

But as a simple illustration, we can easily show that the EGF for D_N would be $G(x)=\frac{e^{-x}}{1-x}\left(1-e^{-0.5x^2}\right)$ – see *Online Encyclopaedia of Integer Sequences* (OEIS) as A158243. Now since $\frac{e^{-x}}{1-x}=1+\frac{x^2}{2}+\frac{x^3}{3}+\frac{3x^4}{8}+\cdots$ and $1-e^{-0.5x^2}=\frac{x^2}{2}-\frac{x^4}{8}+\frac{x^6}{48}+\cdots$ then by definition $G(x)=\frac{x^2}{2}+\frac{x^4}{8}+\frac{x^5}{6}+\cdots=\frac{1}{N!}\sum_{N=0}^{\infty}D_N x^N$. So this implies that $D_2=\frac{2!}{2}=1$, $D_4=\frac{4!}{8}=3$, $D_5=\frac{5!}{6}=20$, etc., which is indeed correct from Table 3 which follows in the Appendix.

Alternatively, to anyone from a digital signal processing (DSP) background $\sum_{N=0}^{\infty}D_N x^N$ above is almost identical to the traditional z-transform for a causal sequence – i.e., $\sum_{N=0}^{\infty}D_N x^{-N}$. And so manipulating the usual inverse z-transform using complex contour integration (see eqn. 4.67 in [2]), rather than expanding the EGF $G(x)$ to get D_N, we can easily show that: $D_N=\frac{N!}{2\pi j}\oint_{|z|=r,\,0<r<1}G(x)x^{-(1+N)}\,dx$. In practice we would now evaluate this contour integration via the '*theorem of residues*' as an alternative method to get D_N (see section 4.5 in [2]).

(ii) We should also note that it may now be argued that (14) (reproduced below) could also have been derived from the well-known *'principle of inclusion/exclusion (PIE)'* used in counting the cardinality of the union of intersecting sets:

$$D_N=\sum_{m=1}^{\lfloor N/2\rfloor}D_N(m)=\sum_{m=1}^{\lfloor N/2\rfloor}(-1)^{m+1}Z_N(m),\quad Z_N(m)=\left(\frac{N!}{2^m m!}\right)\left(\frac{!(N-2m)}{(N-2m)!}\right)$$

But that is definitely with the benefit of hindsight. It was not obvious at the start of this paper to form $Z_N(m)$ in (6), and having done so, again it was not obvious that its method of construction and its effective multiple counting of $D_N(m)$ allowed $Z_N(m)$ to be included in a PIE approach for deriving D_N as above. But most importantly, this paper solves the original problem in more didactic manner, and offers a tutorial style insight to the solution from first principles.

(iii) Now future work might re-examine the icebreaker event where instead of '*pair-swaps*' we might have '*triplet-swaps*'. In the case of a single triplet-swap, attendee A gets the name badge of attendee B who gets the name badge of attendee C who gets the name badge of attendee A. This could similarly invalidate the icebreaker event (as was previously described for pair-swaps) where you could now get stuck in a sort of three person '*limit-cycle*' or alternatively three people may not be called.

More generally, we could also investigate outcomes where we have derangements with a mixture of various r-person swap cycles for different values of $r = 2, 3, 4, \ldots, N-3, N-2, N$ (in this paper we have only considered $r = 2$). These are also called '*derangements with various r-cycles*' in the pure maths literature. Note that a derangement with a full N-cycle will not invalidate the icebreaker event (as every person will eventually get called) but all derangements made up of a mixture of r-cycles $2 \leqslant r \leqslant N-2$ will definitely invalidate the icebreaker event (where $r = N$-1 cannot happen). Note that in Table 1, for $N = 4$ there are nine possible derangements – three with two 2-cycles and six with a single 4-cycle.

(iv) Following on from (iii) we now can define $P_{\text{fail}}(N)$ as the probability that the icebreaker event will fail – i.e., the probability that not everyone will be called. This formula has a much simpler evaluation than that for $P(N)$ with:

$$\begin{aligned} &\text{Probability of guaranteed success} \\ &= 1 - P_{\text{fail}}(N) \\ &= \text{Probability of a derangement containing a single } N\text{-cycle} \end{aligned} \tag{25}$$

Now N objects produce $!N$ derangements, and we can easily show that there is a subset of (N-1)! derangements where each derangement contains a single (but different) N-cycle. Consider as example Table 1 and $N = 4$, where there are (N-1)! = 6 derangements with each having a different N-cycle/4-cycle. So from (25) we can now write:

$$P_{\text{fail}}(N) = 1 - \frac{(N-1)!}{!N} = 1 - \frac{1}{N \sum_{i=0}^{N} (-1)^i / i!} \approx 1 - \frac{e}{N} \quad \text{for large } N \tag{26}$$

N	$P_{\text{fail}}(N)$
2	0
3	0
4	0.333333
5	0.454545
6	0.547170
7	0.611709

Table 3: Probability that the '*icebreaker event*' will fail with not all N attendees being called (see (26)).

It is now clear that for more than five attendees any icebreaker event like this is more likely to fail than succeed – i.e., fail by not calling all N attendees. And so for the 40 attendees who were at my icebreaker event, with$P_{\text{fail}}(40) = 0.9320$ it was almost certain not to succeed.

(v) Finally, it might be interesting to see if these results could also relate to more practical and pressing combinatorial mathematical problems in signal processing, machine learning, communications, physics, genetics, etc. I mean, it is hardly the end of the world if a workshop icebreaker event like this fails because of the statistics of an N-person derangement – or is it?

References

[Wiki] Retrieved from `https://en.wikipedia.org/wiki/Derangement`

[OS89] Oppenheim, A. V. and Schafer, R. W., *Digital Signal Processing*, Prentice-Hall International, 1989.

A Appendix 1

Prove that $\sum_{i=1}^{k}(-1)^{i+1k}C_i = 1, \quad k = 1, 2, 3, \ldots$

Now we know that $(1+x)^k = \sum_{i=0}^{k} {}^kC_i x^i$ and letting x = -1:

$$\Rightarrow \sum_{i=0}^{k} {}^kC_i(-1)^i = 0 \Rightarrow 1 + \sum_{i=1}^{k} {}^kC_i(-1)^i = 0 \Rightarrow \sum_{i=1}^{k}(-1)^{i+1k}C_i = 1$$

B Appendix 2

$\boldsymbol{N}$	Total number of derange-ments, $(!N)$	Number of derangements $(D_N(k))$ with exactly k pair-swaps in each derangement, with $D_N(k)$ from either (5), (10) generalised for any N, or computer simulation	No. of derangements (see (14)) with one or more pair-swaps, $D_N = \sum_{k=1}^{\lfloor N/2 \rfloor} D_N(k)$	Probability, $(P(N))$, of one or more pair-swaps in a derangement, where $P(N) = \frac{D_N}{!N} = \frac{\sum_{k=1}^{\lfloor N/2 \rfloor} D_N(k)}{!N}$	Number of pairs able to swap (see (2)) – i.e., values that $k \geqslant 0$ can take for non-zero $D_N(k)$	Average number of pair-swaps $\frac{}{K(N)}$ in any derange-ment, see (18)
2	1	$D_2(1) = 1$	1	100%	1	1
3	2	$D_3(0) = 2$	0	0%	0	0
4	9	$D_4(0) = 6$ $D_4(2) = 3$	3	33.33%	0,2	0.667
5	44	$D_5(0) = 24$ $D_5(1) = 20$	20	45.45%	0,1	0.455
6	265	$D_6(0) = 160$ $D_6(1) = 90$ $D_6(3) = 15$	105	39.62%	0,1,3	0.509

N	Total number of derange-ments, $(!N)$	Number of derangements $(D_N(k))$ with exactly k pair-swaps in each derangement, with $D_N(k)$ from either (5), (10) generalised for any N, or computer simulation	No. of derangements (see (14)) with one or more pair-swaps, $D_N = \sum_{k=1}^{\lfloor N/2 \rfloor} D_N(k)$	Probability, $(P(N))$, of one or more pair-swaps in a derangement, where $P(N) = \frac{D_N}{!N} = \frac{\sum_{k=1}^{\lfloor N/2 \rfloor} D_N(k)}{!N}$	Number of pairs able to swap (see (2)) – i.e., values that $k \geqslant 0$ can take for non-zero $D_N(k)$	Average number of pair-swaps $\overline{K(N)}$ in any derange-ment, see (18)
7	1854	$D_7(0) = 1130$ $D_7(1) = 504$ $D_7(2) = 210$	714	38.51%	0,1,2	0.493
8	14,833	$D_8(0) = 8988$ $D_8(1) = 4480$ $D_8(2) = 1260$ $D_8(4) = 105$	5845	39.41%	0,1,2,4	0.500
9	133,496	$D_9(0) = 80864$ $D_9(1) = 41040$ $D_9(2) = 9072$ $D_9(3) = 2520$	52632	39.43%	0,1,2,3	0.500

Table 4: Summary of some results for N object derangements for $2 \leqslant N \leqslant 9$. (Interesting to note that the sequence $D_N,\ N = 2, 3, 4, \ldots$ does indeed appear in the *Online Encyclopaedia of Integer Sequences* (OEIS) in A348075 and A158243 but without giving a simple generating expression like (14) in this paper.)