

# Petite énigme de combinatoire

## Énoncé

*Vous faites partie d'un groupe de 500 personnes, rassemblé dans une grande pièce. On donne à chacune de ces 500 personnes un numéro différent entre 1 et 500 (donc il n'y a pas deux personnes ayant le même numéro), mettons que vous recevez le numéro 37, vous êtes le seul à l'avoir, jusqu'ici tout va bien. Dans une autre pièce attenante à celle où votre groupe est rassemblé, il y a 500 casiers. À l'intérieur de chacun de ces 500 casiers est disposé un numéro entre 1 et 500, marqué sur un bout de papier, avec encore une fois aucun doublon : il n'y a pas deux casiers avec le même numéro à l'intérieur. L'on a donc 500 personnes avec toutes un numéro différent d'un côté, et 500 casiers avec tous un numéro différent de l'autre. Tour à tour, chacune des 500 personnes de votre groupe va se rendre dans la pièce aux casiers, et ouvrir puis refermer jusqu'à 250 casiers, pour voir le numéro qui se cache à l'intérieur. S'il voit dans l'un de ces 250 casiers le numéro qui correspond au sien, il gagne, sinon il perd. Dans tous les cas, il retourne dans l'autre pièce, et n'a pas le droit, en aucune manière que ce soit, de communiquer avec les autres personnes du groupe (de sorte qu'en entrant dans la salle des casiers, personne n'ait déjà une idée de quel casier renferme quel numéro). Par exemple, supposons que ce soit votre tour d'y aller : vous ouvrez un premier casier, vous y trouvez, disons, le numéro 412. Ce n'est pas le votre (je vous rappelle que vous avez le numéro 37), vous refermez le casier, vous en ouvrez un second, vous y trouvez le numéro 125. Toujours pas bon, vous refermez, vous en ouvrez un troisième, etc, et ce 250 fois. Si à un moment vous avez trouvé le numéro 37 dans un casier, vous avez gagné, sinon vous avez perdu.*

*On suppose que dans une journée, tout le monde a eu le temps d'aller ouvrir ses 250 casiers. Votre but, à vous les 500 personnes, est que tout le monde ait gagné dans la même journée : il faut que chacune des 500 personnes ait ouvert au moins une fois le casier contenant son numéro, parmi les 250 qu'elle a ouverts. Si ce but est atteint, vous gagnez tous un grand voyage en Australie et le jeu est fini ; sinon, on considère que tout le monde a perdu, et on retente la chance le lendemain, en mélangeant les numéros dans les casiers pendant la nuit (comme ça, encore une fois, personne n'a déjà une idée de quel casier renferme quel numéro en entrant le lendemain dans la salle, même ceux qui ont très bonne mémoire).*

*Vous pensez vous laisser aller au désespoir devant l'impossibilité de la tâche, lorsque l'une des personnes de votre groupe (sûrement un*

*normalien...*) s'exclame : « Si on suit ma méthode, au bout d'une semaine, on aura plus de 9 chances sur 10 d'avoir gagné un voyage en Australie! » *Comment fait-il ?*

## **Eh oui, comment fait-il ?**

### **Le procédé génial**

La stratégie proposée par notre petit génie de normalien est la suivante, qu'il expose en s'ébouriffant les cheveux : les 500 personnes se mettent d'accord pour numéroter toutes de la même façon les 500 casiers, de 1 à 500 encore une fois. Donc on attribue à chaque casier un numéro unique entre 1 et 500. Ensuite, chaque personne, lorsqu'elle ira dans la pièce aux casiers pour ouvrir ses 250 casiers, commencera par se diriger vers le casier qui porte son numéro : si vous avez le numéro 37, vous irez ouvrir en premier le casier numéro 37. Ensuite, vous regardez le numéro qui est à l'intérieur de ce premier casier, et vous irez ouvrir le casier qui porte ce numéro. Et ainsi de suite. Par exemple, si vous trouvez le numéro 412 dans le premier casier, vous allez ouvrir en second le casier numéro 412. Puis, supposons que le papier à l'intérieur de ce casier 412 porte le numéro 125, alors vous irez ouvrir ensuite le casier numéro 125, etc. Ainsi, le papier que vous trouvez à l'intérieur d'un casier détermine le numéro du casier que vous irez ouvrir ensuite. Vous pouvez ouvrir deux fois le même casier par ce système, cela ne pose aucun problème. Lorsque vous avez ouvert 250 casiers (donc pas forcément différents), vous arrêtez ; si, lors de ces ouvertures, vous avez trouvé votre numéro, vous gagnez pour la journée, et sinon, vous perdez, comme toujours. Et comme toujours, vous et vos camarades de groupe partez en Australie si vous gagnez tous la même journée.

Et il se trouve que cette méthode, a priori toute bête, marche très bien ! Grandiose, non ! ? Ce qu'il faut voir, ce qui est fort dans cette méthode, c'est que si vous deviez rouvrir un casier que vous avez déjà ouvert, c'est que vous avez vu votre numéro. En effet, comme vous suivez une genre de « chaîne » de casiers, dont chaque maillon est uniquement déterminé par le maillon précédent (le numéro du prochain casier que vous ouvrez est défini par le numéro que vous trouvez dans le casier précédent), et comme vous avez commencé par le casier qui porte votre numéro, si vous devez ouvrir une seconde fois un casier que vous avez déjà ouvert, alors le premier de ces casiers que vous êtes amenés à rouvrir sera le casier de départ, celui qui porte votre numéro, celui que vous avez ouvert en tout premier (car tous les numéros attribués sont uniques, il n'y a pas deux casiers qui peuvent renfermer un papier avec le même numéro écrit dessus). Et pour être amené à le rouvrir, selon la logique que vous avez suivie, il faut que vous ayez vu le numéro

de ce casier, donc votre numéro, dans un autre casier. Donc vous avez gagné, et les plages d'Australie se rapprochent un peu plus.

### Si c'est si clair, ça doit se prouver...

L'idée se formalise assez bien mathématiquement pour aboutir à la démonstration du résultat annoncé (qui est qu'en une semaine, donc en 7 essais, les 500 personnes ont plus de 90% de chances d'aller faire du surf sur les plages australiennes). Ce qui va suivre utilise des outils de terminale S, des concepts de maths sup' et un raisonnement de maths spé. Ce qui ne veut pas dire que seul un spé peut comprendre, c'est tout à fait à la portée des sup', voire des terminales ; juste que pour y penser et mener à bien le raisonnement *tout seul*, il faut *a priori* être au moins en spé.

Examinons la stratégie décrite précédemment. Ce qu'on fait, simplement, en attribuant des numéros entre 1 et 500 aux casiers, c'est déterminer une permutation de  $[1, 500]$  : à  $k \in [1, 500]$ , on associe l'entier qui est inscrit sur le papier qui se trouve dans le casier numéro  $k$ . Lorsqu'on suit le processus déterminé dans la stratégie, ce que j'ai appelé la « chaîne de casiers », on se contente simplement d'itérer la permutation plusieurs fois. On sait par ailleurs qu'une permutation se décompose en cycles à supports disjoints : on se contente donc, en itérant la permutation, d'en décrire un des cycles. On le parcourt en entier si la longueur de ce cycle est inférieure ou égale à 250, et partiellement si sa longueur est supérieure à 250, car on s'arrête au bout de 250 itérations. Le numéro recherché, celui que l'on porte, celui qui nous renvoie au premier casier ouvert, est ainsi le "dernier" élément rencontré du cycle, si on suppose que celui-ci commence par le numéro trouvé dans le premier casier que l'on a ouvert. Donc si la permutation n'a que des cycles de longueur inférieure ou égale à 250, c'est gagné : chacune des 500 personnes va forcément trouver son numéro lors de sa visite, car va parcourir en entier un cycle dont le dernier élément sera son numéro !

Comme les papiers sont placés au hasard dans les casiers, la permutation est elle-même tirée au hasard, et on aimerait donc calculer la probabilité qu'une permutation de  $[1, 500]$  n'ait que des cycles de longueur inférieure ou égale à 250. Il est plus simple de calculer la probabilité complémentaire : celle qu'une permutation ait au moins un cycle de longueur  $k > 250$ . Notons  $n = 500$  : l'important est juste que  $n$  soit assez grand, et pas qu'il soit précisément égal à 500 (on supposera aussi  $n$  pair pour simplifier les calculs, mais ça ne change fondamentalement rien).

Comptons, avec  $k > \frac{n}{2}$ , le nombre de permutations qui possèdent un cycle de longueur  $k$  : pour en construire une, on commence par

choisir  $k$  éléments parmi les  $n$ , il y a  $\binom{n}{k}$  possibilités ; les  $(n - k)$  restants forment une permutation quelconque<sup>1</sup> de  $[1, n - k]$ , il y a donc  $(n - k)!$  possibilités ; enfin, les  $k$  éléments choisis peuvent s'organiser en  $(k - 1)!$  cycles différents<sup>2</sup>. Ce qui fait finalement  $\binom{n}{k} (n - k)! (k - 1)! = \frac{n!}{k}$  permutations qui possèdent un cycle de longueur  $k$ .

Finalement, la probabilité qu'une permutation ait au moins un cycle de longueur  $k > \frac{n}{2}$  est  $P = \frac{1}{n!} \sum_{k=\frac{n}{2}+1}^n \frac{n!}{k} = \sum_{k=\frac{n}{2}+1}^n \frac{1}{k}$ . Avec

$n = 500$ , la somme est techniquement calculable, mais ça serait pénible, même pour un ordinateur. On peut l'approcher assez efficacement :

on reconnaît la forme  $P = H_n - H_{\frac{n}{2}}$ , où  $H_n = \sum_{k=1}^n \frac{1}{k}$  est le  $n$ ième

nombre harmonique. Comme  $H_n = \ln(n) + \gamma + o(1)$ , on a  $H_n - H_{\frac{n}{2}} = \ln(n) - \ln(\frac{n}{2}) + o(1) = \ln(2) + o(1)$ . En supposant  $n$  assez grand, et 500 est assez grand pour l'approximation que l'on va faire, on obtient  $P \approx \ln(2) = 0,69314\dots$ . On peut noter que  $P$  ne dépend pas vraiment de  $n$ , donc même si on avait 1000 ou 10000 personnes plutôt que 500 (et le nombre idoine de numéros bien sur), la méthode serait toujours aussi efficace.

Finalement, il y a une probabilité égale à  $1 - P \approx 0,30685\dots$  de gagner, en un jour, avec cette stratégie. On est loin de la méthode naïve ! Le calcul montre que la probabilité d'avoir gagné au bout 7 jours, c'est à dire le complémentaire de la probabilité de ne gagner aucun des 7 jours (c'est plus simple à calculer) vaut  $1 - P^7 \approx 1 - (\ln(2))^7 = 0,92131\dots$ . C'est bien supérieur à 90% ! En deux semaines, on passe à 99,4%...

Merci le normalien, et à vous l'Australie !

1. Notons que cela n'est vrai que parce que  $k > \frac{n}{2}$ , sinon l'on devrait soit affiner cette hypothèse soit compter plusieurs fois certaines permutations : si  $k > \frac{n}{2}$ , on est sûr que si on compte une permutation comme possédant un  $k$ -cycle, elle n'aura pas d'autre cycle de longueur supérieure à 250 dans les autres éléments qui la composent, qui peuvent donc être organisés arbitrairement ; ce qui ne serait évidemment pas le cas si on avait  $k \leq \frac{n}{2}$ .

2. En effet, il s'agit là de compter le nombre de permutations circulaires de  $k$  éléments, ou  $k$ -cycles, qui est  $(k - 1)!$  : au premier élément, on peut faire correspondre un des  $k - 1$  restants ; puis à celui-là, on peut faire correspondre un des  $k - 2$  restants, etc. Le dernier s'envoie forcément sur le premier élément, et il y a finalement  $(k - 1)(k - 2) \dots 1 = (k - 1)!$  choix possibles.