RENDEZ-VOUS

VRAIES ET FAUSSES ANOMALIES DU LOTO

Certains tirages, dans les jeux de hasard, ont pu susciter des soupçons de tricherie. Les mathématiques aident à distinguer entre vraies coïncidences statistiques et manipulations frauduleuses.

L'AUTEUR

JEAN-PAUL DELAHAYE
professeur émérite à l'université de Lille et chercheur au laboratoire Cristal (Centre de recherche en informatique, signal et automatique de Lille)

Jean-Paul Delahaye a également publié : **Aux frontières des mathématiques : Kurt Gödel et l'incomplétude** (Dunod, 2025).

Le 1er décembre 2020, le tirage effectué par le loto d'Afrique du Sud (Lotto) fournit la série de nombres 5 – 6 – 7 – 8 – 9 et le numéro complémentaire 10. Vingt personnes ont joué la série complète ainsi que le numéro complémentaire – un nombre anormalement élevé de gagnants pour ce loto. La série sans le numéro complémentaire a par ailleurs été jouée par 70 autres joueurs – là encore, un nombre inhabituellement élevé. Chacun des vingt gagnants principaux empoche l'équivalent de 350000 euros environ, et les 70 autres gagnants chacun 400 euros environ. Le nombre inhabituel de gagnants suscite rapidement des soupçons de tricherie, et déclenche une enquête.

Une grille du loto sud-africain est composée de 5 numéros différents compris entre 1 et 50 et d'un numéro complémentaire compris entre 1 et 20. Il y a donc $50\times49\times48\times47\times46\times20/(5\times4\times3\times2) = 42375200$ grilles possibles. Dans ce décompte, la division par $5\times4\times3\times2$ provient de ce que l'ordre des cinq numéros joués (hors numéro complémentaire) ne compte pas: il ne faut donc pas compter plusieurs fois des séries comportant les mêmes numéros mais placés dans un ordre différent. En France, une grille comporte cinq numéros différents compris entre 1 et 49, et un numéro complémentaire compris entre 1 et 10. Il y a donc $49\times48\times47\times46\times45\times10/(5\times4\times3\times2) = 19068840$ grilles possibles. Le jeu est donc légèrement plus favorable aux joueurs français.

Revenons au tirage 5 – 6 – 7 – 8 – 9 – 10 en Afrique du Sud. Doit-on s'étonner que cette grille soit tombée? Et doit-on s'étonner du nombre exceptionnel de gagnants? Pour un mathématicien, la réponse à ces deux questions est négative.

Il convient d'abord de rappeler que, si le tirage est honnête, toutes les séries sont équiprobables. La série 5 – 6 – 7 – 8 – 9 – 10 est donc tout aussi probable que la série 1 – 3 – 5 – 7 – 9 – 11, la série 6 – 5 – 4 – 3 – 2 – 1 ou la série 2 – 19 – 27 – 35 – 48 – 9. Qu'une série nous apparaisse structurée ou que nous n'y voyons aucune régularité, toutes tombent avec la même probabilité.

Concernant le nombre de gagnants quand une série structurée tombe, il n'y a pas lieu de s'étonner qu'il soit plus élevé qu'à l'accoutumée, car les joueurs jouent souvent de telles séries.

ALÉA MÉCANIQUE OU NUMÉRIQUE

Il faut savoir que, contrairement à bien d'autres jeux de loto, en Afrique du Sud le tirage n'est pas mené avec de vraies balles numérotées, qu'un système mécanique fait sortir une à une d'une urne transparente contenant initialement une bille pour chaque numéro. Depuis 2016, le tirage des numéros gagnants y est opéré par un système électronique (*voir l'encadré 2*) conçu par la société Szrek2Solutions. Le système utilisé a été plusieurs fois expertisé et validé par des autorités indépendantes. Quand il est utilisé, les résultats peuvent être recalculés, et donc contrôlés *a posteriori*.

Helena Pereira, de la société Szrek2Solutions, a expliqué face à des journalistes qui enquêtaient sur l'événement que le tirage du 1er décembre 2020 était parfaitement régulier et ne pouvait pas avoir été manipulé: «La loterie sud-africaine dispose de deux systèmes distincts. Le premier effectue le tirage et le second le vérifie. Cela se fait avant de publier les résultats. Après le tirage il est encore possible de revenir en arrière et de vérifier le tirage.» De fait, les vérifications menées *a posteriori* ont confirmé l'absence d'anomalie: la séquence incriminée est bien le fruit du seul hasard.

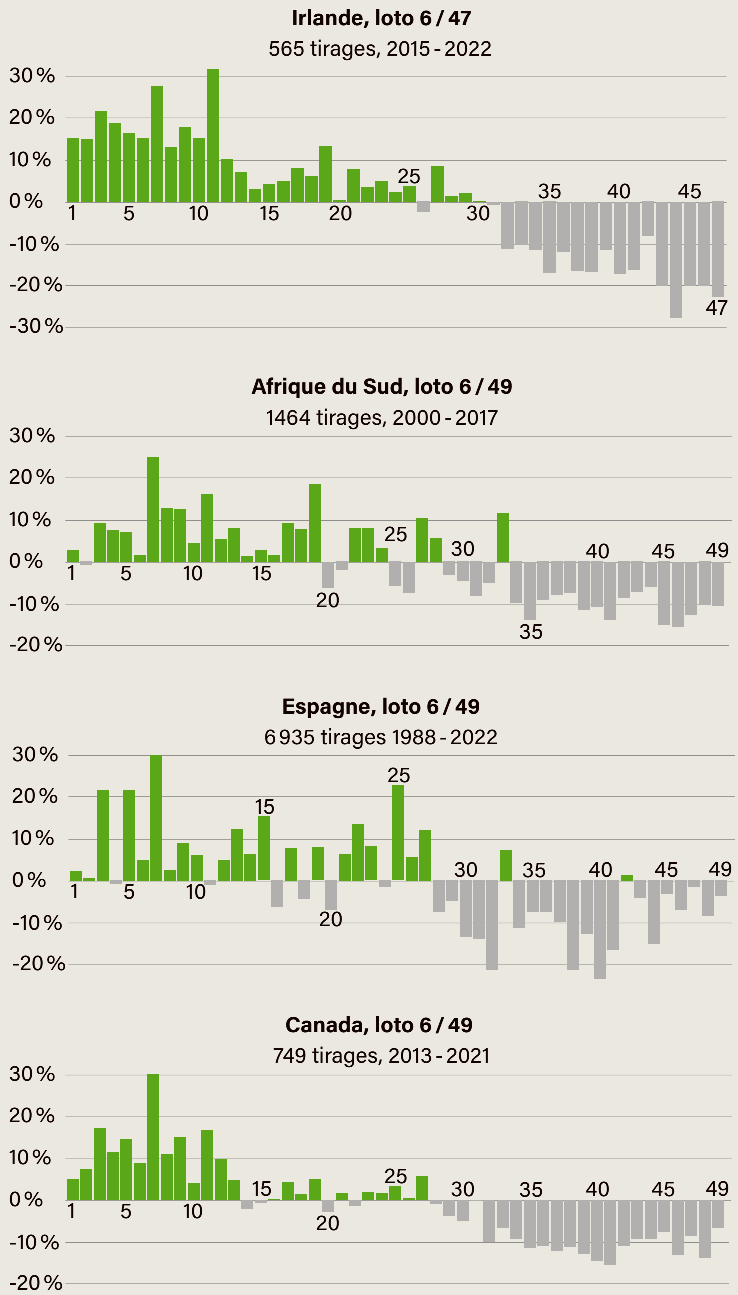
Plus récemment, le 1er octobre 2022, au loto des Philippines, un scandale similaire a agité le monde du jeu. Dans ce loto dit «6/55», les joueurs composent leurs grilles en choisissant 6 nombres différents compris entre 1 et 55. Le nombre total de grilles possibles est donc: $55\times54\times53\times52\times51\times50/(6\times5\times4\times3\times2) = 28\,989\,675$. Dans l'historique de cette loterie, il y a en moyenne moins d'un gagnant par tirage... pourtant, ce 1er octobre 2022, il y a eu un nombre inédit de 433 gagnants! Le gain, équivalent à 4 millions d'euros, a donc été partagé en 433. S'agissait-il d'un hasard extraordinaire, ou d'un trucage impliquant une multitude de joueurs informés à l'avance de la grille qui allait sortir?

Le tirage de ce loto 6/55, contrairement à celui d'Afrique du Sud (et comme c'est toujours le cas en France), utilise des boules numérotées qu'un mécanisme extrait une à une d'une urne en mouvement. Les tirages sont publics et diffusés à la télévision en direct. Cela

D'UN PAYS À L'AUTRE

1

Dans un article récent, les statisticiens James Hanley, de l'université McGill, au Canada, et Michael Cronin, de l'université de Cork, en Irlande, ont déterminé les probabilités avec lesquelles sont joués les numéros des lotos de plusieurs pays. Nous reproduisons ci-dessous quatre de leurs tableaux, qui montrent que les biais des joueurs, tout en variant d'un pays à l'autre, conservent des similitudes assez remarquables. Par exemple, on observe presque toujours une préférence pour le 7, peu d'intérêt pour le 20, et plus généralement une plus faible probabilité de choix des numéros les plus grands.



Préférences des joueurs de loto pour certains numéros, dans différents pays. Pour chaque numéro, le diagramme indique le pourcentage de déviation par rapport à une distribution uniforme (dans laquelle les joueurs choisiraient les nombres de leurs grilles uniformément au hasard).

n'interdit cependant pas d'imaginer des trucages – boules légèrement plus petites ou lestées pour certains numéros, aimants favorisant ou défavorisant certaines boules, etc. Sans prendre en compte les préférences des joueurs pour certains numéros, des calculs de probabilité évaluent qu'un tel nombre de gagnants, parmi les dix millions de joueurs ayant parié ce jour-là, est aussi probable que d'obtenir cent fois de suite « pile » en lançant cent fois de suite une pièce équilibrée. Cela semble donc indiquer une tricherie... mais en réalité, ces calculs omettent un facteur essentiel ! La série incriminée est, en effet, 9 – 45 – 36 – 27 – 18 – 54... où l'on reconnaît les 6 premiers multiples de 9. La grille est donc structurée, comme celle du loto d'Afrique du Sud de 2020 : c'est probablement la meilleure explication du nombre élevé de gagnants. D'ailleurs, on peut noter qu'au même tirage, des centaines de personnes avaient joué la série des multiples de 7 : 7 – 14 – 21 – 28 – 35 – 42.

NUMÉROS PRÉFÉRÉS

Une étude menée par Patrick Roger, chercheur en finance comportementale aujourd'hui en poste à l'EM Strasbourg Business School, a montré que les numéros préférés des joueurs de loto en France sont, par ordre de préférence,

2 ALÉA NUMÉRIQUE ET VÉRIFICATION A POSTERIORI

Un système de tirage au sort numérique peut éviter tout risque de manipulation ou de tricherie... à condition qu'il soit bien conçu. Un critère majeur est la possibilité, pour n'importe qui, de vérifier *a posteriori* que le tirage au sort est bien le fruit du hasard. On détaille ci-dessous un système remplissant ces conditions.

On suppose qu'on dispose d'une bonne fonction de hachage publique *h*, qui à tout fichier numérique *F* associe une empreinte *h*(*F*) – par exemple, une suite de 256 bits. Une telle fonction vérifie les trois propriétés ci-dessous.

(a) *h* est déterministe, mais ce qu'elle produit n'a pas de régularité identifiable : on peut donc assimiler son résultat à un tirage au sort reproductible.

(b) La modification d'un seul caractère du fichier *F* qu'on confie à *h* change totalement le résultat produit.

(c) La fonction *h* n'est pas inversible : en pratique, il est impossible, étant donné une suite *s* de 256 bits, de trouver un fichier *F* tel que *h*(*F*) = *s*, et pour un fichier *G* donné de trouver un autre fichier *G'* tel que *h*(*G'*) = *h*(*G*).

On suppose aussi qu'on dispose d'une fonction *t*, publique, qui associe une grille de loto à toute empreinte calculée par *h*. On suppose de plus que *t* produit des grilles qui sont « bien réparties dans l'ensemble des grilles possibles » : le nombre d'empreintes possibles étant beaucoup plus grand que le nombre de grilles possibles, on exige que, quand on tient compte de tous les résultats que produit *t*, la fonction donne chaque grille possible à peu près le même nombre de fois.

La veille du jour du tirage, dernier jour où l'on peut jouer une grille, l'organisateur de la loterie publie sur une page internet publique la valeur *R* = *h*(*D*), empreinte d'une donnée *D* composée de 256 bits, choisie librement par l'organisateur. Personne ne peut retrouver *D* à partir de *R*, puisque *h* n'est pas inversible.

Le jour du tirage, l'organisateur consulte la première page numérique *J* d'un journal quotidien déterminé en amont, et en calcule l'empreinte *S* = *h*(*J*).

Il calcule ensuite la somme logique de *D* (que lui seul connaît) et de *S* (que tout le monde connaît). Cette somme logique est définie comme la série de bits obtenue en effectuant un « OU exclusif » (noté XOR, en logique) entre les bits de *D* et de *S*, ce qui mélange les deux données. Il en calcule ensuite l'empreinte *X* = *h*(*D* XOR *S*), puis il génère à l'aide de la fonction *t* la grille *G* = *t*(*X*). Il rend enfin publiques la valeur *D* et la grille *G*, qui est la grille gagnante du jour.

Ce procédé prévient toute forme de tricherie.

– L'organisateur ne peut pas tricher en faisant en sorte de connaître la grille en avance ou en la choisissant à sa guise. En effet, quand il a choisi *D*, la première page du journal, *J*, n'était pas connue. Par conséquent, il n'a pas pu manipuler le choix de *G* = *t*(*h*(*D* XOR *h*(*J*)), qui dépend de *J*. Tout le monde peut vérifier que le *D* choisi la veille n'a pas été modifié entre-temps, car en calculant l'empreinte de *D* on doit trouver *R* qui, lui, a été rendu public la veille.

– Le journal ne peut pas tricher, car quand il publie sa première page *J* le jour du tirage, la donnée *D* n'est encore connue que de l'organisateur : le journal ne peut donc pas manipuler la valeur de *G* = *t*(*h*(D XOR *h*(J)), qui dépend de *D*.

- Personne d'autre ne peut tricher, car les paris sont clos avant le jour du tirage, à un moment où *J* n'est pas fixée, et *D* connue seulement de l'organisateur.

Par ailleurs, n'importe qui peut vérifier *a posteriori* qu'il n'y a eu aucune manipulation ou tricherie, car une fois *D* et *G* publiés, chacun peut vérifier que *G* = *t*(*h*(*D* XOR *h*(*J*)).

Veille du tirage

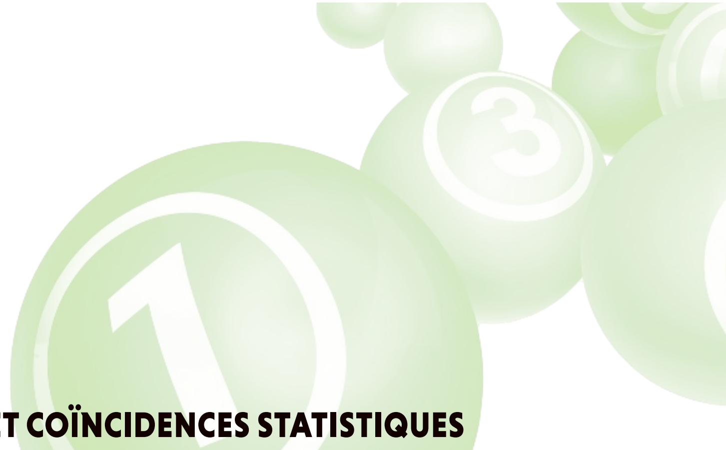
- Les fonctions *h* et *t* sont publiques.
- On choisit publiquement un journal quotidien (qui servira le lendemain).
- L'organisateur choisit une suite *D* de 256 bits.
- Il calcule l'empreinte *R* = *h*(*D*).
- *R* est rendu publique, *D* est gardée secrète.

Jour du tirage

- On appelle *J* la première page du quotidien choisi la veille (et qui donc est publique).
- On calcule l'empreinte *h*(*J*).
- L'organisateur calcule la grille *G* = *t*(*h*(*D* XOR *h*(*J*)))

Après le tirage

- *G* est rendue publique.
- *D* est rendue publique.



3 PARADOXE DES ANNIVERSAIRES ET COÏNCIDENCES STATISTIQUES

On considère une caractéristique individuelle (date d'anniversaire, nombre de cheveux, etc.) qui peut prendre k valeurs qu'on suppose équiprobables pour chaque individu. On réunit n personnes et l'on se demande quelle est la probabilité $P(k, n)$ pour que deux d'entre elles au moins possèdent la même caractéristique.
Il y a k^n combinaisons possibles de ces paramètres, pour une assemblée de n personnes – car il y a k valeurs possibles pour la caractéristique de la première personne, k valeurs possibles pour la caractéristique de la deuxième, et de même jusqu'à la n-ième personne. Parmi ces k^n situations, le nombre de cas où les propriétés des n personnes sont toutes différentes est : $k \times (k-1) \times (k-2) \times \ldots \times (k-(n-1)) = k! / (k-n)!$, car pour déterminer une telle situation, il y a k valeurs possibles pour la caractéristique de la première personne, $(k-1)$ possibilités pour la caractéristique de la deuxième personne, et ainsi de suite jusqu'à la n-ième personne pour laquelle il ne reste que $(k-(n-1))$ valeurs possibles de la caractéristique.
La probabilité que les n personnes aient toute une valeur différente pour la caractéristique considérée est donc : $[k! / (k-n)!] / k^n$. On en déduit immédiatement que la probabilité que deux personnes ou plus aient la même caractéristique est : $P(k, n) = 1 - [k! / (k-n)!] / k^n$.
Dans le cas des anniversaires (en négligeant les années bissextiles), $k = 365$. Si on note $Anni(n) = P(365, n)$ la probabilité que deux personnes dans une assemblée de n individus partagent la même date anniversaire, on obtient les valeurs indiquées dans le tableau ci-contre.

Dès que 23 personnes sont réunies, la probabilité que deux d'entre elles aient la même date anniversaire dépasse 50 % ! Dans une assemblée de 100 personnes, il est presque impossible que toutes les dates anniversaire soient différentes.

n	Anni(n)
2	0,27 %
5	2,71 %
10	11,69 %
20	41,14 %

n	Anni(n)
22	47,57 %
23	50,73 %
50	97,04 %
100	99,999969 %

La formule montre aussi que la probabilité pour que deux personnes dans une assemblée de 2 000 individus possèdent exactement le même nombre de cheveux – nombre qu'on supposera compris entre 0 et 200 000 – est $P(200001, 2000) =$ 99,9956 %. Par ailleurs, dès que l'assemblée comporte au moins 527 personnes, il y a plus de 50 % de chances que deux personnes y aient exactement le même nombre de cheveux : $P(200001, 527) =$ 50,02315 %.
Enfin, pour le loto bulgare de type 6 / 49, où 13 983 816 grilles différentes sont possibles (car chaque grille est composée de 6 numéros différents compris entre 1 et 49), on trouve bien que, dès lors que 4 404 tirages ont été effectués, il y a plus d'une chance sur deux que deux de ces tirages aient donné la même grille gagnante : $P(13983816, 4404) =$ 50,01280 %.

7, 9, 11, 12, 13 et 5, et qu'inversement les numéros les moins joués sont 41, 39, 32, 38, 40 et 3. Le déséquilibre entre les numéros choisis par les joueurs ne doit pas surprendre. Il est dû en particulier – mais pas uniquement – à ce que l'on aime jouer des dates, ce qui favorise les nombres entre 1 et 12 pour les mois, et entre 1 et 31 pour les jours. Il est par ailleurs intéressant de noter que ces préférences sont culturelles, et varient d'un pays à l'autre (*voir l'encadré 1*).

Cette observation a un intérêt pratique pour les joueurs de loto. En effet, avec un tirage honnête, aucun numéro n'est favorisé, aucune grille n'a plus de chances de tomber qu'une autre, et établir une stratégie prenant en compte le passé des tirages est vain, car ce qui a été tiré dans le passé n'a aucune influence sur les futurs tirages. En revanche, les choix des autres joueurs influencent le gain moyen: plus les gagnants sont nombreux, moins leurs gains sont élevés, puisqu'ils se partagent la somme allouée aux vainqueurs. Il faut donc fuir les numéros les plus joués, et préférer les numéros les moins joués. Pour la France on en déduit que jouer la grille 7 – 9 – 11 – 12 – 13 – 5 sera en moyenne très peu rentable, alors que jouer beaucoup de numéros parmi 41, 39, 32, 38, 40 et 43 augmente le gain statistique moyen.

Tâchons de quantifier cette augmentation. Pour simplifier le raisonnement et les calculs, considérons un loto de type 5/50: une grille est composée de 5 numéros différents compris entre 1 et 50, et il n'y a pas de numéro complémentaire. En l'absence de tout biais de la part des joueurs, chacun des 50 numéros serait donc joué avec une probabilité 1/50=2%. Supposons donc, au contraire, que vingt-cinq des cinquante numéros sont favorisés par les joueurs, et qu'ils sont donc chacun joués avec une probabilité de 2,5%. Supposons par ailleurs que les vingt-cinq autres numéros, un peu défavorisés, sont joués chacun avec une probabilité de 1,5%. Notons G le gain moyen, en euros, d'une grille quelconque. Nous allons établir que le gain moyen d'une grille portant des numéros moins souvent joués sera un peu plus élevé que G, car en moyenne il y aura moins de gagnants avec ces grilles que pour une grille quelconque.

Pour faciliter encore le calcul, nous supposons que la probabilité qu'une grille soit jouée est proportionnelle au produit des probabilités de tirage des numéros qu'elle comporte. On suppose aussi que, pour un nombre de parieurs

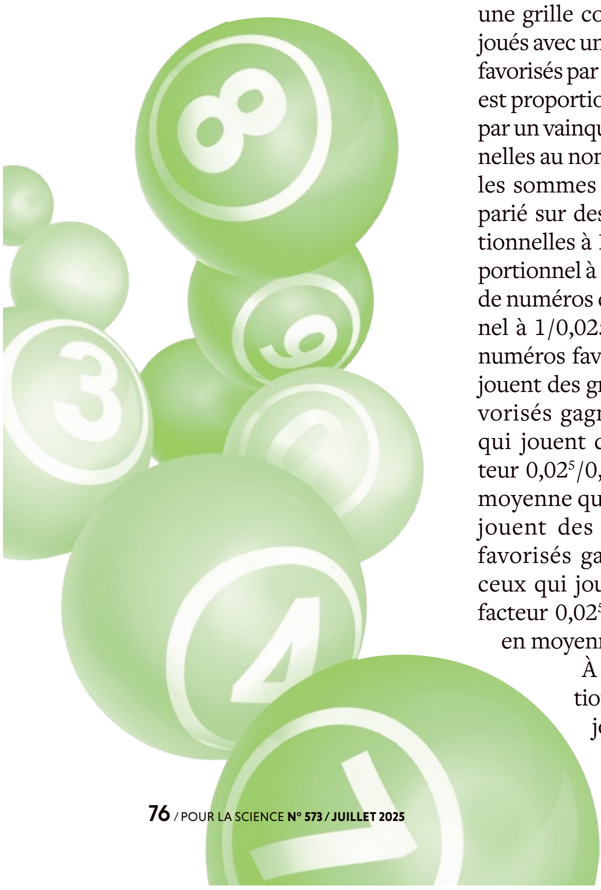
4 UNE ANOMALIE VOLONTAIRE : L'AFFAIRE ROBERT RIBLET

L'histoire des grattages de la Française des jeux (FDJ) est une anecdote remarquable qui illustre combien un système de loterie mal conçu peut permettre à des acteurs bien informés d'en biaiser les résultats. Au milieu des années 2000, Robert Riblet, ingénieur et chef d'entreprise à la retraite, annonce disposer de la preuve que la répartition des gains sur les tickets de grattage de la FDJ n'est pas le seul fait du hasard – contrairement à ce que croient les joueurs et à ce qu'affirme alors l'entreprise. Il s'en explique dans un livre paru en 2008 : « Les tickets de grattage sont vendus dans des paquets qui valent tous 150 euros. Pour un jeu à 1 euro il y a 150 tickets, 75 tickets pour un jeu à 2 euros, 50 tickets pour un jeu à 3 euros. Ce que j'ai découvert, c'est que dans tous ces carnets, il y a toujours un tiers de la valeur totale du carnet qui est remboursée sous forme de petits lots, 1, 2, 3 ou 10 euros maximum. Dans environ trois paquets sur quatre, il y a en plus un lot exceptionnel, qui va de 20 euros jusqu'au très gros lot de 40 000 euros. Problème : dans ces carnets, quand il y a un gros lot, il n'y en a toujours qu'un seul. Autrement dit, une fois qu'il est sorti, il n'y a plus que des misères. Si le détaillant l'a vu sortir, il sait qu'il n'y a plus rien et que les clients qui arrivent après seront perdants, ou au mieux ne gagneront que de tout petits lots. C'est un peu comme si, dans une tombola de kermesse, on continuait à vendre des billets alors que le cadeau à gagner a déjà été emporté. »
Le biais dénoncé est incontestable, d'autant plus que le vendeur des billets qui connaît cette règle de répartition peut en profiter ou en faire profiter ses meilleurs clients : quand il ne reste que très peu de tickets à vendre dans un paquet et qu'aucun gros lot n'a été gagné, alors il devient intéressant d'acheter tout ce qui reste. Conséquence du conflit l'opposant à Robert Riblet, la FDJ fait désormais imprimer sur les tickets la mention : « Au moment de votre achat certains lots ont peut-être déjà été emportés. »

donné, le nombre moyen de gagnants pour une grille est proportionnel à la probabilité que cette grille soit jouée. Cette hypothèse n'est pas absurde: si une grille a deux fois plus de chances d'être jouée, elle sera jouée en moyenne par deux fois plus de joueurs, donc elle rapportera deux fois moins à chacun d'eux.

Il résulte de ces hypothèses raisonnables que le nombre moyen de gagnants pour une grille quelconque est proportionnel à $0{,}02^5$. Pour une grille composée uniquement de numéros joués avec une probabilité de 1,5% (les numéros défavorisés par les joueurs), le nombre de gagnants est cette fois-ci proportionnel (avec la même constante de proportionnalité) à $0{,}015^5$. Et pour une grille composée uniquement de numéros joués avec une probabilité de 2,5% (les numéros favorisés par les joueurs), le nombre de gagnants est proportionnel à $0{,}025^5$. Les sommes gagnées par un vainqueur étant inversement proportionnelles au nombre de gagnants, on en déduit que les sommes gagnées par les vainqueurs ayant parié sur des grilles quelconques sont proportionnelles à $1/0{,}02^5$, que le gain moyen est proportionnel à $1/0{,}015^5$ pour les grilles composées de numéros défavorisés, et qu'il est proportionnel à $1/0{,}025^5$ pour les grilles comportant des numéros favorisés. On en déduit que ceux qui jouent des grilles composées de numéros défavorisés gagnent en moyenne plus que ceux qui jouent des grilles quelconques, d'un facteur $0{,}02^5/0{,}015^5=(4/3)^5=4{,}21$: ils gagnent en moyenne quatre fois plus. De même, ceux qui jouent des grilles composées de numéros favorisés gagnent moins, en moyenne, que ceux qui jouent des grilles quelconques, d'un facteur $0{,}02^5/0{,}025^5=(4/5)^5=0{,}32$: ils gagnent en moyenne trois fois moins.

À la condition d'avoir des informations précises sur les biais de choix des joueurs, il existe donc des stratégies de jeu qui augmentent les gains moyens de manière significative. Notons, d'ailleurs, que la Française des jeux (FDJ) donne des informations (inutiles) sur les numéros tirés dans le passé… mais n'indique rien concernant les numéros préférés des joueurs!



HASARD ET RÉPÉTITION

Un autre type d'événement lié aux tirages des lotos a, par deux fois, provoqué un étonnement mondial et une forte suspicion de tricherie: la répétition, à quelques jours d'intervalle, du même tirage, à une même loterie. Le 4 septembre 2009, le tirage du loto bulgare donne la série 4 – 15 – 23 – 24 – 35 – 42. Six jours plus tard, le tirage se répète. Le 21 septembre 2010, le loto israélien donne la série 13 – 14 – 26 – 32 – 33 – 36. Le 16 octobre de la même année également. Dans les deux cas, ces doublons ont provoqué une levée de boucliers et des soupçons de manipulation. Dans les deux cas, l'organisme chargé du tirage, forcé de s'expliquer, a mené une enquête pour finalement conclure à l'absence de tricherie.

Ces coïncidences sont à mettre en parallèle avec le célèbre «paradoxe des anniversaires»: dans une assemblée de 23 personnes, la probabilité que deux individus aient la même date anniversaire est supérieure à 50%. À première vue, il y a de quoi s'étonner, puisqu'il y a 365 (voire 366) dates anniversaire possibles! Sur un plan statistique, c'est pourtant tout à fait normal, et démontré (*voir l'encadré 3*). En effectuant le même type de calcul dans le cas du loto bulgare, où une grille est composée de 6 numéros différents compris entre 1 et 49 – ce qui donne 13983816 grilles possibles –, on constate que dès que plus de 4404 tirages sont effectués, la probabilité que deux d'entre eux donnent la même série est supérieure à 50%.

Si, pour l'instant, tous les étranges tirages mentionnés ne relèvent pas d'un processus de manipulation frauduleuse, cela ne doit pas conduire à croire qu'il n'y a jamais de tricherie lors des tirages des lotos et loteries. Qu'il s'agisse de dispositifs physiques insuffisamment surveillés ou de procédures électroniques mal conçues, l'histoire montre que des erreurs ont existé, dont les escrocs ont parfois su tirer profit.

On peut mentionner à ce titre la vieille histoire du «Triple Six Fix». Elle concerne un loto américain de l'État de Pennsylvanie, objet en 1980 d'une manipulation par Nick Perry, le présentateur du tirage télévisé. Dans cette loterie, trois machines tirent chacune une balle dans une urne qui en contient 10, numérotées de 1 à 10. Des complices de Nick Perry avaient fabriqué trois séries de fausses balles, numérotées et peintes comme celles qui servaient dans le tirage officiel, et les avaient placées discrètement dans les urnes. Les fausses balles avaient toutes été lestées par un supplément discret de peinture, à l'exception de celles portant les numéros 4 et 6, qui avaient donc plus de chances d'être extraites. Cela rendait les combinaisons avec des 4 et des 6 beaucoup plus probables: les nombres 444, 446, 464, 466, 644, 646, 664 et 666 avaient donc des probabilités de sortie très supérieures aux autres. Le 24 avril, le numéro 666 tombe. Les complices retirent rapidement les balles truquées, les détruisent et remettent en place les balles usuelles. Malheureusement pour les tricheurs, les vendeurs de billets avaient remarqué le nombre anormalement élevé de billets comportant les 8 combinaisons possibles de 4 et de 6, et un coup de téléphone malheureux entre complices a finalement trahi les fraudeurs.

PROGRAMMATION ET MANIPULATION

Cette dernière anecdote est à même de jeter un doute sur les systèmes de tirage mécaniques, qui exigent une fabrication particulièrement soignée puis une surveillance sans faille. Au contraire, les méthodes algorithmiques de tirage avec possibilité de contrôle *a posteriori* fournissent une solution infaillible (*voir l'encadré 2*). Il faut cependant bien noter que cette possibilité de contrôler *a posteriori* les résultats est essentielle, sans quoi les méthodes électroniques sont susceptibles, comme les systèmes mécaniques, d'être altérées pour produire des résultats prévisibles: c'est ce qu'illustre l'histoire de l'informaticien Eddie Tipton.

Le Hot Lotto était une loterie administrée par la Multi-State Lottery Association (MUSL), entre 2002 et 2017. Chaque grille était composée de cinq nombres différents compris entre 1 et 47 et d'un numéro complémentaire compris entre 1 et 19. Le 20 juillet 2015, Eddie Tipton, directeur de la sécurité informatique de la MUSL, a été reconnu coupable de fraude pour avoir faussé un tirage du Hot Lotto en décembre 2010, puis pour avoir tenté de réclamer le prix de manière anonyme. L'enquête approfondie l'obligea à admettre qu'il avait également manipulé, quelques années auparavant, des tirages dans le Colorado, l'Oklahoma, le Wisconsin et le Kansas ainsi que dans l'Iowa. En examinant le programme du générateur de nombres aléatoires utilisé en 2007 pour le tirage du Wisconsin, les experts ont compris que le programme ne donnait des résultats faussés que très rarement: cela ne pouvait se produire qu'à trois dates dans l'année – le 27 mai, le 23 novembre ou le 29 décembre – et à condition que la date soit un mercredi ou un samedi et que le tirage ait lieu après 20 heures.

> **Il existe des stratégies de jeu qui augmentent les gains moyens de manière significative**

Insistons sur le fait que l'utilisation d'un système électronique avec possibilité de contrôle *a posteriori* aurait évité cette tricherie: ce qui fait la sécurité d'un système, ici et comme toujours en sécurité informatique, ne résulte pas du secret dans lequel il est maintenu, mais au contraire de sa bonne conception et de l'accès de tous à sa définition.

Il faut enfin noter que, sans qu'il y ait nécessairement une volonté de frauder la loterie, une mauvaise conception de cette dernière peut induire des biais exploitables par des acteurs bien informés. C'est ce que l'ingénieur Robert Riblet a mis en évidence pour les jeux de grattage de la FDJ (*voir l'encadré 4*): la répartition des gains sur les tickets n'est pas le seul fruit du hasard. Pour ce jeu, non seulement le système de distribution des gains sur les tickets est injuste par certains aspects, mais surtout les buralistes vendant les tickets peuvent anticiper, en partie, les gains de leurs clients, et les en informer... voire profiter eux-mêmes de cette information privilégiée.

Vrais trucages, fausses anomalies, biais culturels, déséquilibres volontaires ou non: ces histoires montrent qu'une bonne compréhension mathématique des décomptes, des calculs probabilistes soigneux et une évaluation logique attentive sont les seuls moyens de distinguer ce qui est normal de ce qui indique des tricheries. ■

BIBLIOGRAPHIE

A. Kavousi *et al.*, SoK : Public randomness, *IEEE 9th European Symposium on Security and Privacy*, 2024.

J. Hanley et M. Cronin, Which lottery numbers do players play ?, *Significance*, 2023.

T. Wang *et al.*, Number preferences in lotteries, *Judgment and Decision Making*, 2016.

A. Lenstra et B. Wesolowski, A random zoo : Sloth, unicorn, and trx, *Cryptology ePrint Archive*, 2015.

G. Delbos et R. Riblet, *100 % des perdants ont tenté leur chance, l'affaire des jeux de grattage*, Seuil, 2008.

P. Roger et M. H. Broihanne, Efficiency of betting markets and rationality of players : Evidence from the French 6 / 49 Lotto, *Journal of Applied Statistics*, 2007.

P. Roger, *Lotomania, approche scientifique du jeu et du comportement des joueurs*, Village Mondial, 2005.