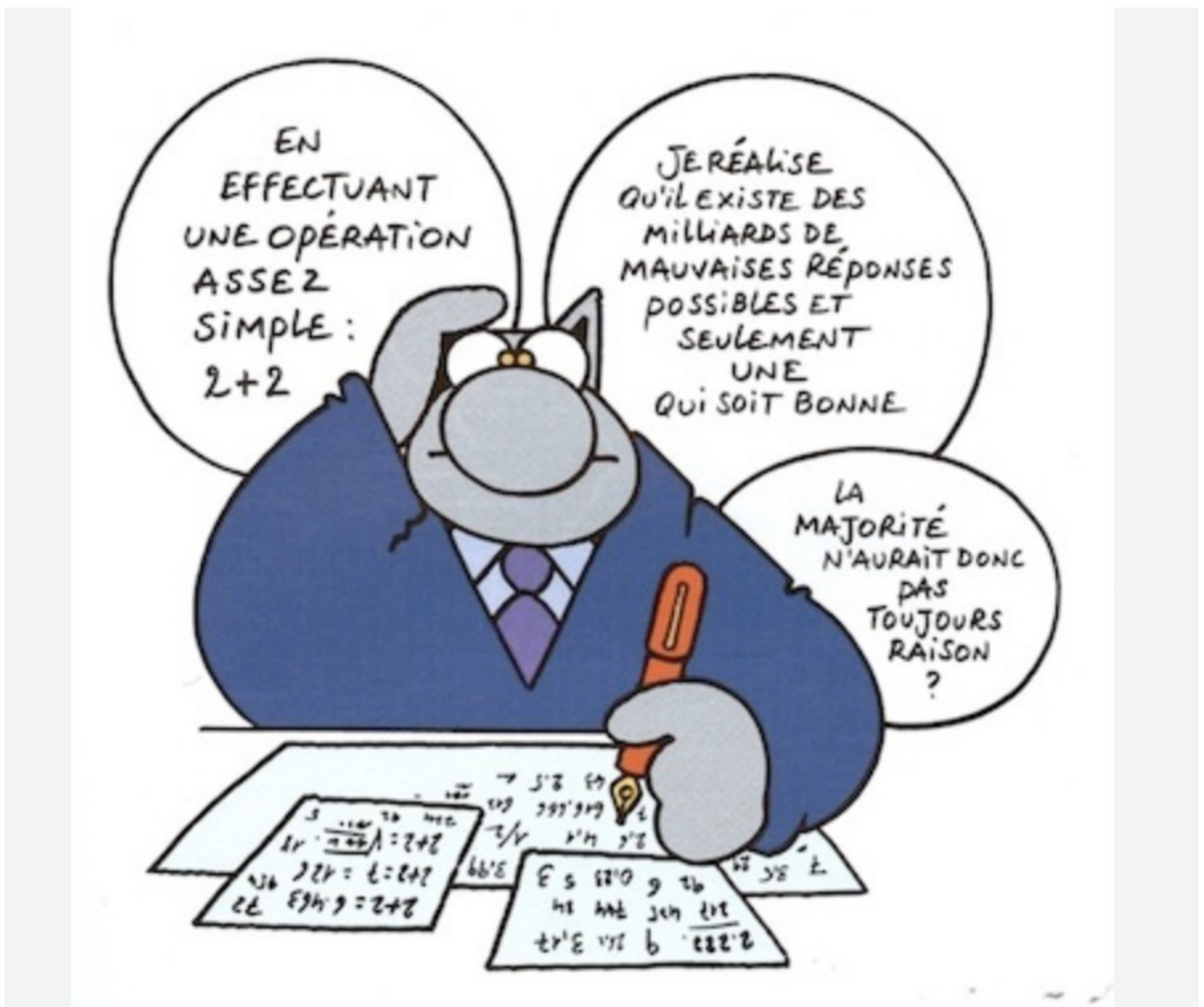


On joue aux probabilités ? Connaissez-vous le paradoxe des anniversaires ?

écrit par Alain Falento | 12 octobre 2023





Les probabilités de la vie quotidienne se caractérisent par le fait que tout le monde peut comprendre le problème.

Néanmoins, il est très facile de faire une mauvaise analyse, mais beaucoup moins de comprendre son erreur.

Rappelons qu'une probabilité est un nombre compris entre 0 et 1, qui mesure le degré de plausibilité d'un événement, avec une probabilité de 0 pour un événement impossible (obtenir 7 en lançant un dé à 6 faces) jusqu'à l'événement certain (tirer pile ou face en lançant une pièce de monnaie).

Sans que nous en ayons forcément conscience, l'intégralité des activités humaines modernes sont construites autour de cette notion (choix d'un traitement médical, assurances,

contrôle qualité des produits, etc.).

Même la Justice demande aux jurés de rendre leur verdict au-delà de tout « doute raisonnable », alors que personne ne s'est donné la peine de quantifier ce que cela signifiait.



À notre échelle individuelle, nous évaluons constamment des

probabilités (que nous appelons « chances ») pour prendre des décisions telles que la filière scolaire des enfants, le modèle de voiture que nous achetons ou la compagnie aérienne que nous empruntons.

Cependant, il est étonnant de constater à quel point notre évaluation peut être éloignée de la réalité, y compris pour des problèmes extrêmement simples.

Le célèbre *paradoxe des anniversaires* – qui n'a d'ailleurs de paradoxe que le nom – en est un exemple.

La question posée est la suivante : *si vous réunissez 23 personnes dans une pièce, quelle est la probabilité qu'au moins d'eux d'entre elles aient la même date d'anniversaire (jour de l'année) ?*

La résolution de ce petit problème ne présente aucune difficulté pour un élève de terminale (démonstration en annexe à la fin de l'article), mais tout l'intérêt réside dans la comparaison entre l'évaluation rapide que nous en faisons, et la réalité du résultat.

Lorsque cette question fut posée à votre serviteur (qui étudiait pourtant la chose depuis plus de quinze ans), son réflexe immédiat fut de diviser 23 par 365 pour avoir une évaluation, et sa réponse fut « environ une chance sur 10 ».

Ce raisonnement n'a aucun sens, je le conçois, et ferait rougir de honte tous mes anciens professeurs de mathématiques.

Tous ceux qui ont accepté de se prêter à ce test honnêtement (agrégés de mathématiques, docteurs en sciences, étudiants, plombiers, employés de bureau, menuisiers ou agriculteurs) ont tous sans exception donné une réponse similaire, comprise en 1 et 15%.

Or, la réponse est 51%, soit plus d'une chance sur deux...

Plus incroyable encore, un rapide calcul montre que la probabilité monte à 70% pour 30 personnes et plus de 99% pour 57 personnes.

Pourtant, il ne semble pas bien difficile de faire asseoir 57 personnes sur 365 chaises sans que deux d'entre elles choisissent la même chaise.

C'est apparemment beaucoup plus difficile si chacune d'elles choisit un numéro de chaise au hasard.

Les mathématiques (dans la vraie vie...)



Comment expliquer cette différence entre l'évaluation a priori et la réalité ? Et surtout pour quelle raison ceux qui ont un niveau en mathématiques probabilistes supérieur aux autres ne semblent nullement avantagés ?

Je n'ai pas de réponse définitive à ces deux questions, mais il existe un faisceau d'explications, se situant à l'intersection de la biologie et des sciences cognitives.

Le premier obstacle entre nous et la solution est tout d'abord de nature linguistique : nous nous faisons une mauvaise représentation mentale de la question posée.

À l'instar des référendums, on nous pose une question et nous répondons à une autre.

« Au moins deux personnes » ne signifie pas exactement deux personnes, mais deux ou plus, et pas forcément toutes à la même date.

Il faut prendre mentalement en considération tous les cas tels que 2 personnes le 3 février, ou 2 personnes les 31 mars et 3 personnes le 12 avril.

Il est déjà impossible de visualiser mentalement toutes les distributions possibles du cas simple de deux personnes sur 5 jours, alors les 23 personnes sur 365 jours sont hors de portée d'un cerveau humain.

De plus, il apparaît que le cerveau humain est par nature « proportionnel ».

Autant il nous est évident que si nous doublons notre vitesse sur un trajet, nous allons mettre deux fois moins de temps pour le parcourir, autant les phénomènes suivant des lois exponentielles, comme la démographie par exemple, sont contre intuitifs.

Par exemple, la distance de freinage d'une voiture dépendant du carré de la vitesse, il suffira d'augmenter la vitesse de 41% pour doubler cette distance. Le moins que l'on puisse dire, c'est que cela ne saute pas aux yeux.

Dans le cas qui nous occupe, bien évidemment, plus on ajoute de personnes, et plus le nombre de distributions possibles

des dates d'anniversaire explose, atteignant des chiffres incommensurables ($8.56 \cdot 10^{58}$ possibilités pour 23 personnes).

Le nombre de personnes augmentant, il est évidemment de plus de plus difficile qu'aucune d'entre elles n'ait la même date d'anniversaire, mais il nous est impossible de l'évaluer mentalement, car les nombres en jeu dépassent ce que nous sommes en mesure de nous représenter.

Quant au calcul de la probabilité, qui correspond à l'évaluation du ratio entre ces deux nombres, il est bien entendu inaccessible au commun des mortels.

Il semble que le cerveau de l'être humain ne soit pas « câblé » pour évaluer correctement des probabilités ; notre cerveau n'a apparemment pas évolué en adéquation avec les sciences que nous avons développées.

C'est d'ailleurs probablement également la raison pour laquelle il nous est impossible de nous représenter les concepts de mécanique quantique, et que les plus grands spécialistes, comme le Prix Nobel de Physique *Alain Aspect*, avouent en toute honnêteté se contenter du formalisme mathématique.

Quant à savoir pour quelle raison ceux qui ont étudié les mathématiques n'ont aucun avantage compétitif dans la résolution de ce type de problème, la réponse réside dans le fait que leur cerveau reste celui d'un être humain ordinaire.

Un mathématicien aura certes la capacité d'analyser et de résoudre le problème, mais son premier réflexe reptilien sera de répondre la même chose que tout le monde.

Avant de recevoir un charmant message de la part d'un lecteur, me demandant à quoi servent « ces finasseries sans intérêt » (sic), et que tout le monde « se fout royalement

de savoir quel est le résultat d'un problème comme celui-ci » (sic), nous prendrons comme exemple un problème similaire qui fera apparaître toute l'importance de cette famille de cas.

Depuis l'avènement de l'ADN, utilisé pour la première fois par la police scientifique britannique en 1986, des bases de données, certaines contenant les profils de plusieurs centaines de milliers d'individus ont été générées à travers le monde, et particulièrement aux États-Unis.

Ces profils contiennent 13 loci par individu (un locus étant un emplacement sur un chromosome).

La question s'est donc très vite posée de déterminer si ces 13 *loci* étaient suffisants pour identifier à coup sûr, et de façon unique, un individu dans une base de données, et surtout ce qu'il allait advenir de la fiabilité de l'identification lorsque certains *loci* sont détériorés sur un prélèvement.

On se retrouve ainsi sur un problème de la même nature que celui qu'évoque cet article.

Il est même clairement évident que certains accusés ont été libérés indument, ou accusés à tort, suite à une analyse de probabilité erronée délivrée par un avocat astucieux devant un juge incapable d'en comprendre le contenu.

Ainsi, certaines finasseries probabilistes menées correctement auraient pu éviter de funestes erreurs judiciaires.

Alain Falento

Démonstration :

Soit A l'événement : Il y a au moins 2 personnes ayant la même date d'anniversaire.

Soit CA l'événement complémentaire de A : les 23 personnes ont toutes une date d'anniversaire différente.

Soient P(A), P(CA) les probabilités respectives des événements A et CA :

$$P(CA) = (365 \times 364 \times 363 \times \dots \times (365 - 23 + 1)) / 365^{23}$$

Donc $P(A) = 1 - P(CA)$

La probabilité de A est d'ainsi d'environ 51%.