



HAL
open science

A la recherche des lois de probabilités de D'Alembert

Léo Gerville-Réache

► **To cite this version:**

| Léo Gerville-Réache. A la recherche des lois de probabilités de D'Alembert. 2017. hal-01534662

HAL Id: hal-01534662

<https://hal.science/hal-01534662>

Preprint submitted on 7 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A LA RECHERCHE DES LOIS DE PROBABILITE DE D'ALEMBERT

Léo Gerville-Réache - Univ. Bordeaux, CNRS, IMS, UMR 5218, 33405 Talence, France

Résumé. Les positions de D'Alembert en matière de probabilités ont d'avantage fait l'objet de sévères et rapides critiques que de tentatives d'analyses profondes. Bertrand disait: "*L'esprit de D'Alembert, habituellement juste et fin, déraisonnait complètement sur le Calcul des probabilités*". Bernoulli aurait même qualifié ses réflexions de "*ridicules*". Aujourd'hui encore, le raisonnement de D'Alembert sur le jeu croix ou pile est repris dans l'enseignement des probabilités pour montrer "ce qu'il ne faut pas faire" et cultive par la même son discrédit sur le sujet. Ses

doutes et questions et sa recherche non-aboutie sur les lois de probabilité des événements mêlés et non mêlés, le jeu croix ou pile et le paradoxe de St Petersburg ont bien marqué l'histoire des probabilités mais y ont essentiellement stagné. Et si ce grand savant avait une autre vision des probabilités, une vision épistémique au moins aussi rationnelle que celle des "*calculateurs*" de l'époque, qu'il n'est pas parvenu à mettre en relation avec la règle de succession de Laplace?

1) Introduction

Jean le Rond D'Alembert (1717-1783), l'un des plus prestigieux savants du 18^{ème} siècle, est évidemment connu pour avoir dirigé avec Denis Diderot la rédaction d'une encyclopédie mythique. Il est également connu mais de manière moins positive pour ses positions atypiques, voir uniques, sur quelques problèmes de probabilités. En particulier, son article analysant le jeu croix ou pile, publié en 1754 a fait l'objet d'une critique quasi unanime, le plus souvent feutrée (due à son rang) mais parfois virulente en particulier de la part de Daniel Bernoulli.

Il ne faudrait pas croire que la contribution de D'Alembert aux probabilités se résume à cela. Même si celui-ci considérait que la question des probabilités était mineure, il y consacra clairement du temps et de l'énergie afin d'exposer le plus précisément et pédagogiquement possibles ses doutes et questions sur les principes de l'époque : le dénombrement des combinaisons, l'équiprobabilité, le calcul de l'espérance mathématique, les prémisses des probabilités conditionnelles... Principalement, et son écrit de 1780 (trois ans avant sa mort) dans son mémoire 52 en atteste, D'Alembert chercha "désespérément" la (ou les) loi(s) de probabilité qui traduirait sa pensée immuable. Il y en a deux précisément. L'une traduirait pourquoi il convient de parier 2 contre 1 et non pas 3 contre 1 (comme toute le monde l'affirme) dans le jeu croix ou pile et l'autre montrerait pourquoi une suite non mêlée de croix serait, bien que *mathématiquement* équiprobable, *physiquement* moins probable qu'une suite mêlée de croix et de piles.

Pour ne reprendre que la question des suites mêlées et commencer à doter notre propos d'un certain intérêt, regardons un instant la question suivante : lançons deux fois (de manière indépendante) une pièce de l'époque équilibrée et demandons la probabilité que le résultat soit pile-croix ainsi que la probabilité que le résultat soit pile-pile. Tout le monde s'accordera pour dire que les deux résultats ont la même probabilité de se produire : 1/4. On est également tenté de dire que ces deux événements sont, d'un point de vue probabiliste,

similaires ; équivalents. Regardons maintenant le problème suivant : quelle est la probabilité que pile-croix apparaissent en au plus trois lancers indépendants et qu'en est-il de pile-pile ?

Pour pile-pile, les issues favorables sont *PP* et *CPP*. La probabilité recherchée est donc $1/4+1/8=3/8$. Pour pile-croix, les issues favorables sont *PC*, *PPC* et *CPC*. La probabilité recherchée est donc $1/4+1/8+1/8=1/2$. Ce phénomène est connu sous le nom de paradoxe de Walter Penney (1969). Nous savons également que le nombre moyen de lancers pour l'apparition de *PC* est de 4 alors qu'il est de 6 pour *PP*. Fort de cela, pouvons continuer à dire que *PP* et *PC* sont des combinaisons équivalentes d'un point de vu probabiliste?

Etait-ce ce genre de paradoxe, ou plus justement de surprise probabiliste qui généra chez D'Alembert ses doutes les plus profonds? Cela est difficile à dire puisqu'aucun de ses écrits n'en fait mention à notre connaissance. Mais cela nous a convaincu de partir à la recherche de la loi de probabilité de D'Alembert. Certes l'aventure était risquée et incertaine mais il nous a semblé nécessaire de donner toutes ses chances aux doutes de ce grand homme de sciences dont il est difficile de croire qu'il fut autant dans l'erreur que l'on a pu le lire.

A notre grande surprise, ce n'est pas une mais bien deux lois de probabilité que D'Alembert cherchait. Deux lois apparemment contradictoires qui plus est, augmentant encore l'effet paradoxal de ses propos.

2) L'emblème de la discorde : le jeu croix ou pile

L'article CROIX OU PILE, publié dans le tome 1 de Mathématiques de l'Encyclopédie Méthodique en 1754, a fait couler beaucoup d'encre. Quasiment tous ses contemporains expliquaient la grossière erreur de D'Alembert, les critiques étant à la fois nombreuses et virulentes. Pourtant, comme nous allons le voir maintenant, en développant l'un de nos propos précédents (voir Léo Gerville-Réache 2017), D'Alembert avait en tête une première loi de probabilité bien plus cohérente qu'il n'y paraît...

CROIX OU PILE, (analyse des hasards).

Ce jeu, qui est très connu, & qui n'a pas besoin de définition, nous fournira les réflexions suivantes. On demande combien il y a à parier qu'on amènera croix en jouant deux coups consécutifs. La réponse qu'on trouvera dans les auteurs, & suivant les principes ordinaires, est celle-ci.

Il y a quatre combinaisons :

<i>Premier coup :</i>	<i>Deuxième coup :</i>
<i>Croix</i>	<i>Croix</i>
<i>Pile</i>	<i>Croix</i>
<i>Croix</i>	<i>Pile</i>
<i>Pile</i>	<i>Pile</i>

De ces quatre combinaisons, une seule fait perdre & trois font gagner ; il y a donc 3 contre 1 à parier en faveur du joueur qui jette la pièce. S'il pariait en trois coups, on trouveroit huit combinaisons, dont une seule fait perdre, & sept font gagner ; ainsi, il y aurait 7 contre 1 à parier.

Cependant cela est-il bien exact ? Car, pour ne prendre ici que le cas de deux coups, ne faut-il pas réduire à une les deux combinaisons qui donnent croix au premier coup ? Car, dès qu'une fois croix est venu, le jeu est fini, & le second coup est compté pour rien.

Ainsi, il n'y a proprement que trois combinaisons de possibles :

*Croix, premier coup
Pile, croix, premier & second coup
Pile, pile, premier & second coup.
Donc il n'y a que 2 contre 1 à parier.*

De même, dans le cas de trois coups, on trouvera :

*Croix
Pile, croix
Pile, pile, croix
Pile, pile, pile*

Donc il n'y a que 3 contre 1 à parier.

Ceci est digne, ce me semble, de l'attention des calculateurs, & irait à réformer bien des règles unanimement reçues sur les jeux de hasard.

En bon "élève", Pierre-Simon de Laplace (et bien d'autres) tente d'expliquer à D'Alembert son erreur et le bien fondé des réponses "3 contre 1" et "7 contre 1". Par l'analyse de ce jeu, D'Alembert provoque les penseurs de l'époque (voir Michel Paty 1988) :

L'hypothèse d'équiprobabilité, qui est l'objet central de ses critiques, ne lui paraît pas physiquement fondée, précisément parce qu'elle s'abstrait, à ses yeux, totalement et sans justification de ces caractères, causal et temporel, des événements physiques réels"

On pourrait être tenté de considérer le point de vue de D'Alembert comme une interprétation causaliste et chronologiste de la probabilité conditionnelle.

Problème que l'on rencontre souvent chez les élèves et étudiants (voir par exemple André Totohasina 1994). Mais il n'en est rien. D'Alembert ne parle pas de théories mathématiques mais d'événements physiques réels pour lesquels une modélisation probabiliste adéquate ne pourrait s'abstraire des causes et du temps. Il semble cependant que son analyse fasse, en réalité, également l'hypothèse de l'équiprobabilité (celle des observables). Mais ce n'est pas exactement cela, comme il le précise en 1761 dans ses Opuscules Mathématiques tome 2, pages 21-22 :

Je ne voudrais pas cependant regarder en toute rigueur les trois coups dont il s'agit, comme également possibles. Car 1°. il pourroit se faire en effet (& je suis même porté à le croire), que le cas pile croix ne fût pas exactement aussi probable que le cas croix seul ; mais le rapport des probabilités me paroît inappréciable.

Aussi, c'est à défaut de pouvoir déterminer précisément les probabilités auxquelles il croit, que D'Alembert préférerait l'équiprobabilité des combinaisons observables à l'équiprobabilité des combinaisons théoriquement possibles.

Nous savons maintenant ce que nous cherchons. Nous cherchons ce rapport de probabilité entre pile-croix et croix de telle sorte que dans le jeu croix ou pile en deux coups, il y ait 2 contre 1 à parier qu'on amènera croix, 3 contre 1 en trois coups... n contre 1 en n coups.

Dans son essai philosophique sur les probabilités, Pierre-Simon de Laplace en 1774 propose une règle d'attribution (on dirait aujourd'hui *d'estimation*) de la probabilité d'un événement en fonction des observations passées. Supposons qu'un événement ait seulement deux tirages possibles valant "succès" et "échec". Avec l'hypothèse que l'on sache peu ou rien a priori en rapport aux probabilités relatives aux tirages, Laplace propose une formule de probabilité pour que le tirage suivant soit un succès :

$$P(\text{le prochain tirage soit un succès}) = (s+1)/(n+2),$$

où s est le nombre de succès observés précédemment et n est le nombre total des essais observés¹.

En fait, Laplace, père du principe d'indifférence (ou de la raison insuffisante), érige ce principe, non pas en principe d'attribution d'une probabilité a priori en l'absence de toute autre information, mais comme probabilité de départ pour l'application d'un maximum de vraisemblance "bayésien". Pour être précis, la règle de succession de Laplace s'obtient par un a priori bayésien uniforme sur la probabilité de l'événement considéré. Par exemple, en mettant une loi uniforme continue sur $[0;1]$ sur la probabilité de croix, alors l'estimation de $P(C)$ après avoir effectué n lancers

¹ La règle de succession de Laplace est en fait une méthode d'estimation séquentielle de la probabilité d'un succès. En statistique, on utilise généralement l'estimateur du maximum de vraisemblance qui est : $P(\text{le prochain tirage soit un succès}) = s/n$.

Dès que n est supérieur à 0, cette estimation est calculable. En revanche, pour $s=0$, l'estimation donnera 0 quel que soit $n>0$. Aussi cette estimation peut s'avérer bien peu perspicace.

indépendants et ayant observé n piles consécutifs sera :

$$P(\text{le prochain tirage soit croix}) = 1/(n+2),$$

Soit... Mais dans quelle mesure ce principe d'estimation viendrait au secours de l'analyse de D'Alembert?

3) La première loi de D'Alembert

On peut regretter que Laplace (ni aucun autre) n'ait cherché dans cette direction une solution à la question de D'Alembert. Pourtant les deux hommes se connaissent fort bien et se respectent (c'est d'ailleurs D'Alembert qui recrutera Laplace). En effet, Laplace analyse le jeu comme ses contemporains et rejoint ainsi la pensée unique de l'époque. Il n'avait pas vu dans les propos de D'Alembert l'analogie avec sa propre règle de succession...

Nous avons vu que D'Alembert remet essentiellement en cause l'équiprobabilité. Aussi, même lorsqu'on lance une pièce, il est possible que $P(C)$ et $P(P)$ diffèrent. Pour autant, tant qu'il n'existe aucune raison (ou méthode) pour attribuer des probabilités différentes, on ne peut que se plier à l'indifférence et à l'équiprobabilité. En revanche, toute observation pourrait venir modifier ces probabilités (ces estimations) du premier jour. Dans l'analyse du jeu croix ou pile, que se passe-t-il en appliquant la règle de Laplace?

La probabilité estimée, par la première loi de D'Alembert (P_{A1}), que la pièce tombe sur croix au premier lancer est $P_{A1}(C_1)=1/2$. Il en est de même pour $P_{A1}(P_1)$. Au deuxième lancer, on estimera :

$$P_{A1}(P_1 \wedge C_2) = P_{A1}(C_2|P_1) * P_{A1}(P_1) = 1/3 * 1/2 = 1/6,$$

$$P_{A1}(P_1 \wedge P_2) = P_{A1}(P_2|P_1) * P_{A1}(P_1) = 2/3 * 1/2 = 1/3.$$

Aussi, les probabilités estimées sont les suivantes: croix: 1/2, pile-croix: 1/6 et pile-pile: 1/3. On en déduit alors que la probabilité d'amener croix en deux coups est de 2/3 et qu'il y a donc "2 contre 1" à parier. Coïncidence? Non, car cela se poursuit pour le jeu en trois coups :

$$P_{A1}(P_1 \wedge P_2 \wedge C_3) = P_{A1}(C_3|P_1 \wedge P_2) * P_{A1}(P_1 \wedge P_2) = 1/12,$$

$$P_{A1}(P_1 \wedge P_2 \wedge P_3) = P_{A1}(P_3|P_1 \wedge P_2) * P_{A1}(P_1 \wedge P_2) = 1/4,$$

Il est alors facile de voir que :

$$P_{A1}(C_1) + P_{A1}(P_1 \wedge C_2) + P_{A1}(P_1 \wedge P_2 \wedge C_3) = 3/4.$$

Et ainsi de suite on montre qu'il y a n contre 1 à parier au jeu croix ou pile en n coups et pourtant les $n+1$ événements observables ne sont pas équiprobables (ou plus précisément équi-estimés). La cohérence est totale puisque par exemple :

$$P_{A1}(P_1) = P_{A1}(P_1 \wedge C_2) + P_{A1}(P_1 \wedge P_2) = P_{A1}(C_1) \text{ et}$$

$$P_{A1}(P_1 \wedge P_2) = P_{A1}(P_1 \wedge P_2 \wedge C_3) + P_{A1}(P_1 \wedge P_2 \wedge P_3).$$

A ce stade de notre analyse, on est en droit de rétorquer que l'on met dans la détermination des probabilités du jeu des estimations basées sur des hypothèses et non sur des faits (distinction cependant essentielle pour

D'Alembert). En particulier, on dit que si la pièce ne tombe pas sur croix au premier lancer, on devra réviser la probabilité de croix au deuxième lancer. Soit... mais avant que le jeu ne commence, cela n'est qu'une hypothèse. Alors pourquoi ne pas faire les calculs avec une probabilité constante de 1/2? Pour répondre à cette objection visiblement fondée, proposons une petite variante du jeu qui décompose les paris :

- Jouons en un coup;
- il y a donc à parier 1 pour 1 que croix sorte;
- pile sort; on rejoue;
- combien y a-t-il maintenant à parier sur croix?

Si l'on applique la règle de succession de Laplace, on estime $P(C_2|P_1)=1/3$ donc je dois maintenant parier 1 contre 2. Ainsi, chaque pari est ici estimé à somme nulle. Le groupement des deux paris est également à somme nulle. Mais que donne-t-il?

- Au premier lancer, je gagne 1 ou je perds 1 avec une probabilité estimée de 1/2;
- si je gagne le jeu s'arrête et je gagne 1.
- si je perds (cela avec une probabilité estimée de 1/2) alors je perds 1 et je parie encore sur croix;
- si je gagne, je gagne 2 et si je perds, je perds 1.

L'espérance estimée totale du jeu est nulle:

$$1/2 * 1 + 1/2 * [1/3 * (2) + 2/3 * (-1) - 1] = 2/3 * 1 + 1/3 * (-2) = 0.$$

Maintenant en pariant à 2 contre 1 sur un jeu en deux coups, je gagne avec probabilité 2/3 et je perds avec probabilité 1/3. En pariant à 2 contre 1, mon espérance de gain vaut $2/3 * 1 + 1/3 * (-2) = 0$.

Dans le jeu, parier séquentiellement 1 contre 1 sur croix en un coup puis si je perds, 1 contre 2 sur croix est exactement équivalent à parier à 2 contre 1 sur croix en deux coups. Les gains et les pertes en fin de jeu ainsi que les probabilités associées sont identiques. Il est alors bien délicat ici (à moins de remettre en cause le principe de réflexion de Van Fraassen) d'imaginer que les paris puissent globalement différer simplement par le fait que les paris sont simultanés ou séquentiels.

La proposition de pari de D'Alembert n'est pas fautive, elle est même statistiquement plus cohérente que celle de Bernoulli (qui est cependant logique). On peut ne pas y adhérer mais c'est alors en réalité la loi de succession de Laplace que l'on refuse. On refuse cette estimation bayésienne pourtant non informative et basée sur le principe d'indifférence. Ce refus est un choix, pas une vérité mathématique et encore moins une vérité logique. En fait, il semble que D'Alembert veut prendre en compte le conflit intérieur qui se produirait nécessairement à chaque nouvelle apparition de pile (i.e. la pièce pourrait ne pas être équilibrée...). En effet, il veut intégrer dans son pari, la possibilité que les probabilités effectives (objectives) de croix et pile puissent ne pas être égales. C'est ce que Bernoulli ne suppose pas. Au fur et à mesure que les piles apparaissent, soit on continue à croire que pile et croix sont équiprobables, soit on ré-estime les probabilités.

Soit... mais que dit la répétition du jeu? Il est clair que si l'on répète le jeu, la fréquence de la combinaison pile-croix sera de 1/4 et pas de 1/6 (Voir par exemple Parzys 2007). Aussi, bien que l'on trouve une explication à l'analyse de D'Alembert, celui-ci est dans l'erreur! C'est un argument que l'on retrouve souvent dans l'analyse des modèles probabilistes. Mais on fait ici une erreur grossière. Il est clair que si l'on suppose que les issues pile et croix sont équiprobables, une simulation donnera des fréquences proches pour pile et croix et si pile sort au premier lancer, la fréquence de croix au deuxième lancer sera proche de celle de pile (avec des lancers indépendants).

La probabilité pile-croix égale à 1/6 correspond au cas où la probabilité de croix suit une loi uniforme réelle sur [0;1]. C'est la justification dont nous avons déjà parlé pour la règle de succession de Laplace. Si nous devons faire une simulation, nous devons alors générer un nombre pseudo-aléatoire X uniforme entre 0 et 1 puis un deuxième Y indépendamment. Si $Y < X$ alors la pièce est tombée sur croix sinon sur pile. Si l'issue est pile, nous générons alors un troisième nombre pseudo-aléatoire Z , indépendant des deux premiers et comparons Z à X . Si $Z < X$ alors l'issue du deuxième lancer sera croix et sinon pile. La première fréquence qui nous intéresse est celle des piles au premier tirage: $P(Y > X)$. Cette probabilité vaut 1/2 et la fréquence dans la simulation est bien proche de 1/2. La deuxième fréquence qui nous intéresse est celle des croix au deuxième tirage lorsque pile est tombée au premier tirage : $P(Z < X | Y > X)$. Cette probabilité vaut 1/3 et la fréquence dans la simulation est bien proche de 1/3. La fréquence de pile-croix dans la simulation est alors bien proche de 1/6. Aussi, l'estimation de la probabilité de pile-croix n'est pas subjective. Elle correspond bien à une réalité statistique: la réalité d'une pièce dont la probabilité de croix est la réalisation inconnue d'une variable aléatoire réelle uniforme sur [0;1].

Les doutes de D'Alembert semblent trouver une raison épistémique cohérente, bien loin de l'équiprobabilité apparente que ses paris nous laissent croire. La loi du "croix ou pile" propose d'attribuer la valeur 1/6 à la probabilité de pile-croix (et classiquement 1/2 à la probabilité de croix). Le rapport entre ces deux probabilités, que D'Alembert pensait indéfinissable, serait finalement de un à trois. D'Alembert avait bien raison! Raison d'avoir des doutes sur les sommes à parier dans le jeu croix ou pile. Mais, lorsque deux raisonnements aboutissent à des conclusions différentes, on pense qu'il y a soit une erreur soit un *paradoxe*. Ici, il n'en est rien. Les deux conclusions (3 contre 1 et 2 contre 1) proviennent de deux interprétations raisonnables du principe d'indifférence : l'une "mathématique" et l'autre... "physique".

Par exemple la série *PCCPPP* aura une probabilité de :

$$P_{AI}(PCCPPP) = 1/2 * 1/3 * 2/4 * 2/5 * 3/6 * 4/7 = 1/105.$$

Enfin, soit X_i l'historique des i premiers lancers et n_i le

nombre de croix obtenu jusque-là, la première loi de D'Alembert, suivant la règle de Laplace, dit:

$$P_{AI}(C_{i+1}|X_i) = P_{AI}(C_{i+1}|n_i) = (1+n_i)/(2+i),$$

$$P_{AI}(P_{i+1}|X_i) = P_{AI}(P_{i+1}|n_i) = 1 - (1+n_i)/(2+i),$$

Il écrivait pour introduire son propos "Doutes et questions sur le calcul des probabilités" en 1767 :

On se plaint assez communément que les formules des mathématiques, appliquées aux objets de la nature, ne se trouvent que trop en défaut. Personne néanmoins n'avait encore aperçu ou cru apercevoir cet inconvénient dans le calcul des probabilités. J'ai osé le premier proposer des doutes sur quelques principes qui servent de base à ce calcul. De grands géomètres ont jugé ces doutes dignes d'attention; autres grands géomètres les ont trouvés absurdes; car pourquoi adoucerais-je les termes dont ils se sont servis? La question est de savoir s'ils ont eu tort de les employer, et en ce cas ils auraient doublement tort.

Cette recherche n'est pourtant pas terminée. En effet, la première loi de D'Alembert dit par exemple que la probabilité de pile-pile-pile vaut 1/4 et celle de pile-pile-croix, 1/12. Aussi, pile-pile-pile serait trois fois plus probable que pile-pile-croix. Or D'Alembert semblait penser précisément le contraire. En effet, dans son questionnement sur les probabilités des séries mêlées et non mêlées et ses réflexions sur le jeu de Saint Pétersbourg, il affirme qu'il est physiquement plus probable que l'on observe pile-pile-croix que pile-pile-pile (comme le sous-entend, d'une certaine manière, le paradoxe de Penney). Mais alors, D'Alembert qui pourrait bien être réhabilité dans son analyse du jeu croix ou pile serait en réalité en contradiction complète avec lui-même, critiquant tout et son contraire? Ce serait une fois encore ne pas se pencher suffisamment sur ses doutes décidément *justes et fins*.

Dans sa réflexion sur le jeu de Saint Pétersbourg et les probabilités de la nature, il écrit toujours en 1767 :

Ce qu'on vient de dire est fondé sur la supposition que pile ne soit pas arrivé de suite un très-grand nombre de fois : car il serait plus probable que c'est l'effet de quelque cause particulière dans la construction de la pièce, et pour lors il y aurait de l'avantage à parier que pile arriverait encore. Quoi qu'il en soit, j'imagine qu'il n'y a point de joueur sage qui ne doive dans ce cas être embarrassé pour savoir s'il pariera croix ou pile, tandis qu'au commencement du jeu, il dira, sans hésiter croix ou pile indifféremment.

Je demande donc en conséquence :

1°. *Si parmi les différentes combinaisons qu'un jeu peut admettre, on ne doit pas exclure celles où le même effet arriverait un grand nombre de fois de suite, au moins lorsqu'on voudra appliquer le calcul à la nature?*

2°. *Supposons qu'on doive exclure les combinaisons où le même effet arrivera, par exemple, 20 fois de suite; sur quel pied envisagera-t-on les combinaisons où le même effet arrivera 19 fois, 18 fois, etc.? Il me*

paraît peu conséquent de les regarder comme aussi possibles que celles où les effets seraient mêlés. Car s'il est aussi possible, par exemple, que croix arrive 19 fois de suite, qu'il l'est que pile arrive au premier coup, croix ensuite, ensuite pile deux fois si l'on veut, et aussi du reste, en mêlant croix et pile ensemble sans faire arriver long-temps de suite l'un ou l'autre; je demande pourquoi on exclurait absolument, comme ne devant jamais arriver dans la nature, le cas où croix viendrait vingt fois de suite? Comment se pourrait-il que pile pût arriver 19 fois de suite, aussi bien que tout autre coup, et que pile ne pût arriver 20 fois de suite?

Pour moi, je ne vois à cela qu'une réponse raisonnable: c'est que la probabilité d'une combinaison où le même effet est supposé arriver plusieurs fois de suite, est d'autant plus petite, toutes choses d'ailleurs égales, que ce nombre de fois est plus grand, en sorte que quand il est très-grand, la probabilité est absolument nulle ou comme nulle, et que quand il est assez petit, la probabilité n'est que peu ou point diminuée par cette considération.

D'assigner la loi de cette diminution, c'est ce que ni moi, ni personne, je crois, ne peut faire: mais je pense en avoir assez dit pour convaincre mes lecteurs que les principes du calcul des probabilités pourraient bien avoir besoin de quelques restrictions lorsqu'on voudra les envisager physiquement.

Ici, D'Alembert raisonne en supposant que l'on connaît les probabilités de croix et pile et que celles-ci valent $1/2$. Il est alors dans le cadre du problème probabiliste de Penney, de la probabilité et du temps d'apparition d'une suite donnée au cours de lancers successifs indépendants.

Si les issues Croix et Pile sont équiprobables, alors selon D'Alembert, la probabilité de PP serait moindre que celle de PC ! Ce qui est certain c'est que lorsque l'on réalise des lancers indépendants d'une pièce équilibrée, le temps d'attente jusqu'à la première apparition de PP est de 6 alors que celui de PC n'est que de 4. Cela signifie bien d'une certaine manière que les deux événements ne sont pas identiquement équiprobables... Mais que nous dit Penney de plus?

La probabilité d'obtenir PP comme PC en deux coups vaut $1/4$ mais la probabilité d'obtenir PP en au plus trois coups vaut $3/8$ alors qu'elle vaut $4/8$ pour PC . Aussi, dans un jeu en trois coups maximum, il y a 3 à parier contre 5 sur PP alors qu'il y a 1 à parier contre 1 sur PC . Dans ce dernier jeu, PC est plus probable que PP . S'il l'on augmente le nombre de coups possibles, la supériorité de PC persiste (même si à l'infini, les deux probabilités se rejoignent évidemment). La question est donc la suivante : dans quelle mesure doit-on prendre en compte cet état de fait? Quelle probabilité doit-on définir pour un événement dont la probabilité d'apparition est toujours supérieure ou égale à celle d'un autre événement mais qui à l'infini n'est pas plus fréquent?

4) La crédibilité selon D'Alembert

En supposant que les probabilités de croix et pile soient chacune de $1/2$ et les lancers indépendants, cet ensemble d'hypothèses serait d'avantage confirmé par l'apparition d'une suite mêlée que non-mêlée. Dans sa réflexion, D'Alembert part maintenant de l'hypothèse que les probabilités de croix et pile sont égales et les lancers indépendants. Il en conclut qu'une série mélangée (on pourrait dire complexe) confirme d'avantage les hypothèses qu'une série non mêlée (on pourrait dire simple). Cela doit nous rappeler le concept de test statistique et plus précisément celui de la p-value qui mesure la crédibilité d'une hypothèse.

Aujourd'hui, pour tester l'hypothèse qu'une pièce est équilibrée, on calcule une p-value basée sur une loi binomiale. Précisément, sur n lancers indépendants d'une pièce dont s lancers sont tombés sur croix (ou face de nos jours), le calcul de la p-value associée à l'hypothèse que la pièce est équilibrée est obtenu via les probabilités d'une loi binomiale de paramètres n et $1/2$. Cette p-value mesure la crédibilité que les résultats des lancers accordent à l'hypothèse d'équiprobabilité. Pour autant, nous savons bien que cette p-value réduit la problématique dans le sens où, par exemple, les séries $PPPPCCCCCC$, $PCPCPCPCPC$ ou encore $PCCCPCPCPC$ ont la même p-value (l'ordre d'apparition des côtés de la pièce ne rentre pas en ligne de compte). En regardant ces trois séries de près, on se rend compte que les deux premières séries semblent très "symétriques"; "moins aléatoires"; "moins complexes"; les lancers ne semblent pas indépendants. Bien sûr, il existe un(des) test(s) statistique(s) qui, via une p-value, analyse(nt) la dépendance (certaines formes de dépendance). Mais nous savons bien que toutes ces p-values sont réductrices et qu'en réalité c'est la question extrêmement délicate de la définition même d'une suite aléatoire qui est sous-jacente.

D'Alembert cherche donc une mesure de la crédibilité de l'hypothèse conjointe d'équiprobabilité et d'indépendance. C'est cette crédibilité physique qui doit être, selon lui, la probabilité que l'on doit attribuer aux différentes combinaisons possibles. L'hypothèse conjointe d'équiprobabilité et d'indépendance est d'avantage confirmée par certaines combinaisons que par d'autres. C'est cette sorte de degré de confirmation qui serait alors la probabilité physique qu'il conviendrait d'attribuer à chaque combinaison. D'Alembert raisonne par induction et cherche finalement une de réciproque à sa première loi de probabilité.

Comme il ne trouva pas le lien entre sa première loi de probabilité et la règle de succession de Laplace, il ne trouvera pas cette réciproque. Pourtant il tentera de l'expliciter bien d'avantage que la première. Dans son mémoire 52 de 1780, les points 18 à 49 sont consacrés à cette tentative (malheureusement non finalisée). Voici les points qui nous semblent essentiels :

18. Voyons maintenant si on ne trouveroit pas un résultat plus conforme à la vérité, en supposant que tous les cas ne soient pas également possibles, & que lorsque pile, par exemple, est arrivé une ou plusieurs fois de suite, il y a lieu d'espérer que croix viendra ensuite, plutôt que pile, au moins si la pièce n'a pas plus de penchant à tomber d'un coté que de l'autre.²

19. Supposons donc que, si pile est arrivé au premier coup, la probabilité que croix arrivera au second soit $(1+a)/2$, au lieu de $1/2$, a étant une quantité très-petite; que si pile est arrivé les deux premiers coups, la probabilité que croix arrivera au troisième coup est $(1+a+b)/2$, et ainsi de suite, de manière que $a+b+c+d, \&c.$ ne soit jamais $=1$, afin que la probabilité que ne devienne pas certitude absolue. Cela posé,

20. La probabilité que croix arrivera au premier coup, est $1/2$.

21. Celle qu'il n'arrivera qu'au second coup, est le produit de $1/2$, la probabilité que pile arrivera au premier coup, par $(1+a)/2$, la probabilité que croix arrivera en ce cas au second coup.

22. La probabilité que pile arrivera encore au second coup est $(1-a)/2$, & par conséquent la probabilité qu'il arrivera deux coups de suite est $1/2*(1-a)/2$; d'où la probabilité que croix n'arrivera qu'au troisième coup, est $1/2*(1-a)/2*(1+a+b)/2$.

...

49. Dans la théorie ordinaire, & en parlant de tous les principes admis jusqu'à présent par les Géomètres sur l'indifférence des événements semblables successifs ou non successifs, lorsqu'un événement, par exemple, croix, est arrivé plusieurs fois de suite, & qui avoit d'ailleurs aucune raison de croire qu'il dût arriver ainsi, il est clair qu'il y a quelque probabilité que croix avoit plus de penchant à venir que pile. Mais comment estimer cette probabilité après la chute successive & supposée de croix un certain nombre de fois de suite? Cette question a rapport à la recherche de la probabilité des causes par les événements dont plusieurs savant Géomètres se sont occupés. Voyez dans les Transactions philosophiques de 1763 & 1764, les recherches de MM. Bayes et Price sur ce sujet, & celle de M. de la Place dans le VI^e Volume des Mémoires des Savans Etrangers, présentés à l'Académie des Sciences, & imprimés en 1774.⁴

D'Alembert développe dans cet écrit les principes de l'estimation des probabilités des combinaisons mêlées et non-mêlées. Nous avons déjà précisé que nous cherchons en réalité une mesure de la "crédibilité" des hypothèses conjointes d'équiprobabilité et d'indépendances dans une suite de croix et pile. D'Alembert cherche ici un degré rationnel de croyance.

² Au lieu de "également possible" on dirait "également crédible" au regard de la notion de p-value et du développement précédant. Il est essentiel de noter également que D'Alembert fait bien l'hypothèse d'équiprobabilité mais également (en sous-entendu) l'hypothèse d'indépendance entre les lancers.

³ Les points 23 à 48 sont consacrés à l'analyse des propriétés de cohérence nécessaires aux valeurs a, b, c, \dots . Il reprend également le

5) La deuxième loi de D'Alembert

Au lieu de noter a, b, c, \dots les constantes de D'Alembert, nous les noterons a_2, a_3, a_4, \dots . Avec cette notation, l'indice de a correspond au nombre de lancers de la série.

Aussi, en un lancer, on a :

$$P(\text{croix})=P(\text{pile})=1/2,$$

en deux lancers:

$$P(\text{pile-croix})=1/2*(1+a_2)/2 \text{ et}$$

$$P(\text{pile-pile})=1/2*(1-a_2)/2,$$

en trois lancers:

$$P(\text{pile-pile-croix})=1/2*(1-a_2)/2*(1+a_2+a_3)/2 \text{ et}$$

$$P(\text{pile-pile-pile})=1/2*(1-a_2)/2*(1-a_2-a_3)/2,$$

Les constantes a_i tant recherchées par D'Alembert sont les suivantes :

$$a_2=1/(1+2), a_3=1/(1+2+3), \dots a_i = 1/\sum_{k=1}^i k.$$

En trois lancers par exemple, on a:

$$P(\text{pile-pile-croix})=1/2*(1-1/3)/2*(1+1/3+1/6)/2=1/8 \text{ et}$$

$$P(\text{pile-pile-pile})=1/2*(1-1/3)/2*(1-1/3-1/6)/2=1/24,$$

D'où proviennent ces constantes a_i ?

Nous avons vu précédemment que la règle de succession de Laplace est un principe d'estimation bayésienne avec un a priori uniforme sur l'intervalle réel $[0;1]$ de la probabilité de croix (ou de pile). Elle dit que si nous ne savons rien des probabilités de pile et croix, lorsque pile est par exemple obtenu au premier lancer, la probabilité que pile sorte au deuxième lancer doit être estimée à $(1+1)/(2+1)=2/3$ et que la probabilité que croix sorte à ce même deuxième lancer doit être estimée à $1/(2+1)=1/3$. On part de l'hypothèse que la loi de probabilité de croix est inconnue et que des lancers indépendants nous renseignera sur cette probabilité. Quelle peut-être la «réciproque» de ce principe?

Si pile sort au premier lancer et que premièrement, les probabilités de pile et croix sont identiques et deuxièmement, que les lancers sont indépendants, alors la probabilité (le degré rationnel de croyance) que croix sorte au deuxième lancer est le complémentaire à 1 de l'estimation donnée par la règle de succession de Laplace. Précisément, en notant P_{A1} première loi de D'Alembert et P_{A2} la réciproque de cette loi (deuxième loi de D'Alembert), on aurait par exemple:

$$P_{A2}(C_2|P_1)=1- P_{A1}(C_2|P_1)=1-1/(2+1)=2/3.$$

jeu de Saint-Pétersbourg pour en conclure que la mise rationnelle est nécessairement finie. Pour autant, il ne parvient pas à établir ces valeurs salvatrices.

⁴ Ce dernier point est incroyable puisqu'il cite à la fois Bayes et Laplace qui sont les constituants de sa première loi de probabilité et également, comme nous allons maintenant le voir, les constituants de la deuxième.

Cela signifie ici que la croyance que croix vienne après pile en deux lancers est deux fois plus grande que la croyance que pile vienne après pile.

D'Alembert dans la description de sa deuxième loi ne s'intéresse qu'aux événements où il y a une succession de piles puis, soit pile soit croix pour finir. Sa formalisation en terme de constante a, b, c, \dots fonctionne dans ce cadre. En revanche pour des combinaisons quelconques, cette formulation n'est pas adaptée. En revanche, la formalisation en termes d'estimation "selon Laplace" reste valide pour toute combinaison.

Soit X_i , l'historique des i premiers lancers et n_i le nombre de croix obtenu jusque-là, on a :

$$P_{A2}(C_{i+1}|X_i) = P_{A2}(C_{i+1}|n_i) = 1 - (1+n_i)/(2+i) = 1 - P_{A1}(C_{i+1}|X_i),$$

$$P_{A2}(P_{i+1}|X_i) = P_{A2}(P_{i+1}|n_i) = (1+n_i)/(2+i) = 1 - P_{A1}(P_{i+1}|X_i),$$

d'où on obtient par exemple :

$$P_{A2}(P_1 \wedge C_2 \wedge P_3 \wedge P_4) = 1/2 * 2/3 * 2/4 * 2/5 = 1/15,$$

$$P_{A2}(P_1 \wedge C_2 \wedge P_3 \wedge C_4) = 1/2 * 2/3 * 2/4 * 3/5 = 1/10,$$

La deuxième loi de probabilité de D'Alembert est-elle en contradiction avec la première? Non...

Pour argumenter notre réponse, il est d'abord important de noter que ces deux lois compensent l'équiprobabilité. En effet, sous l'hypothèse d'équiprobabilité et d'indépendance, on a par exemple :

$$(P_{A1}(C_2|P_1) + P_{A2}(C_2|P_1))/2 = P(C_2|P_1) = 1/2.$$

Aussi, la première loi de D'Alembert donne une probabilité de l'événement $C_2|P_1$ en dessous de 1/2 et la deuxième loi, une probabilité au-dessus, de telle sorte que la moyenne de ces deux probabilités est de 1/2. P_{A2} est donc la loi «compensatrice» de P_{A1} . Cette propriété définit d'ailleurs de manière unique P_{A2} , dans le sens où c'est la seule loi possédant cette propriété. Plus généralement, on a :

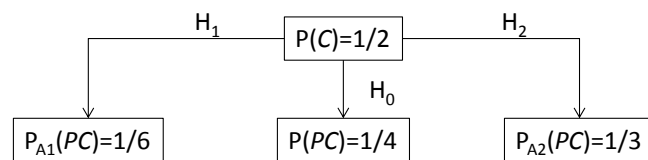
$$(P_{A1}(C_{i+1}|X_i) + P_{A2}(C_{i+1}|X_i))/2 = 1/2.$$

On est néanmoins tenté de croire que ces deux lois sont contradictoires dans la mesure où elles attribuent à un même événement des probabilités différentes. Certes, cela semble paradoxal mais ce n'est pas ainsi qu'il convient de raisonner. En effet, chaque loi de probabilité est issue d'une modélisation spécifique qui se fonde sur des hypothèses.

La question est alors de savoir quelles seraient les bonnes hypothèses. Comme nous l'avons vu plus haut, il n'y a pas de contradiction entre l'analyse "standard" du jeu croix ou pile de Bernoulli et la première loi de D'Alembert.⁵

⁵ La première loi de D'Alembert est même d'avantage en accord avec la réalité des connaissances des qualités physiques de la pièce. En effet, l'analyse standard (Bernoulli) utilise le principe d'indifférence pour faire l'hypothèse que les deux issues de la pièce sont équiprobables et ne remet pas en question cette hypothèse. Pourtant, les réalisations de lancers successifs apportent de l'information sur la réalité des probabilités de la pièce qui devrait, maintenant que

La confrontation entre la première et la deuxième loi de D'Alembert est du même ressort: ces deux lois attribuent des probabilités à un même objet sous deux angles de vue différents (des hypothèses différentes). La divergence des probabilités calculées n'est pas un critère de contradiction. Ce que nous dit D'Alembert c'est que la manière d'attribuer les probabilités aux événements de la nature n'est pas nécessairement unique et que l'on peut envisager plusieurs voies difficilement hiérarchisables. C'est la grande question du choix de la modélisation d'un phénomène réel qui est en jeu.



Probabilités selon les hypothèses.

Avec H_0 on est sous l'hypothèse d'indépendance et d'équiprobabilité des lancers et on en déduit que la probabilité de PC vaut 1/4. Avec H_1 , on est sous l'hypothèse que la probabilité que la pièce tombe sur croix est la réalisation inconnue d'une variable aléatoire uniforme réelle sur $[0;1]$ et on en déduit que la probabilité de PC vaut 1/6. Avec H_2 , on est sous l'hypothèse d'une possible équiprobabilité et indépendance des lancers. On attribue alors à chaque combinaison possible une probabilité qui caractérise le degré rationnel de croyance⁶ que la combinaison confère à H_2 . La probabilité de PC est alors de 1/3.

6) Conclusion

Il serait difficile de conclure cette recherche sans revenir sur l'une des motivations principales de D'Alembert à l'origine de ses recherches de la loi de probabilités des séries mêlées et non-mêlées. Il s'agit du célèbre jeu Saint Pétersbourg. Toujours dans son papier "Doutes et questions sur le calcul des probabilités" de 1767, il écrit :

Pierre joue avec Paul à croix ou pile, avec cette condition que si Paul amène pile au premier coup, il donnera un écu à Pierre; s'il n'amène pile qu'au deuxième coup, deux écus; s'il n'amène pile qu'au troisième coup, quatre écus; au quatrième, huit écus; au cinquième, seize; et ainsi de suite jusqu'à ce que pile vienne; on demande l'espérance de Paul, ou ce qui est la même chose, ce qu'il doit demander à Pierre avant que le jeu commence, pour jouer avec lui à jeu égal, ou, comme on l'exprime d'ordinaire, pour son enjeu.

D'Alembert trouve inadmissible la solution

notre connaissance a augmenté, être prise en compte pour l'attribution des probabilités du futur lancer.

⁶ Ce degré n'est pas sans défaut. Il confère par exemple une crédibilité maximale à la suite alternée PCPCPC... Pour autant, il pourrait participer à la délicate réflexion sur la construction d'une mesure calculable de la complexité aléatoire d'une suite binaire...

mathématique de ce jeu qui consiste à dire que l'espérance du jeu est infinie et que Paul doit demander un somme "infinie" à Pierre pour jouer. Il ajoute plus loin:

Aucun mathématicien ne conteste cette conséquence; mais il n'en ait aucun qui ne sente et n'avoue que le résultat en est absurde, et qu'il n'y a pas de joueur qui voulût à un pareil jeu risquer seulement cinquante écus, et même beaucoup moins. Plusieurs grands mathématiciens se sont efforcés de résoudre ce cas singulier. Mais leurs solutions, qui ne s'accordent nullement, et qui sont tirées de circonstances étrangères à la question, prouvent seulement combien la question est embarrassante.

D'Alembert s'interroge sur la pertinence même de l'espérance morale comme critère de décision⁷. Mais sa priorité est l'idée qu'il serait physiquement bien moins probable que pile sorte aussi tardivement que le calcul mathématique des probabilités le dicte.

Nicolas de Béguelin, l'un des rares "avocats" de D'Alembert, donnait, un an plus tard, une lecture à l'Académie des Sciences "sur l'usage du principe de raison suffisante dans le calcul des probabilités". Il l'introduisait par ces quelques mots :

*J'ai montré dans un Mémoire précédant que la doctrine des probabilités étoit uniquement fondée sur le principe de la raison suffisante; il ne seroit donc pas surprenant que les Mathématiciens ne fussent pas d'accord entr'eux dans la solution des problèmes qui ont la probabilité pour objet; leurs calculs sont de vérités nécessaires, mais la nature du sujet auquel ils les appliquent ne l'est pas. Les vérités contingentes ne peuvent être démontrées qu'en partant d'une supposition; & quelque plausible qu'une supposition soit, elle n'en exclut pas nécessairement d'autres, qui peuvent servir de base à d'autres calculs, & donner par conséquent des résultats différents.*⁸

Le raisonnement de Béguelin le conduisait à établir la mise rationnelle de Pierre à environ 4,9 écus.⁹

Plus empiriquement, Haim Levy et Marshall Sarnat en 1984 ont indiqué dans leur manuel, malheureusement sans donner de détails sur leur procédure, qu'ils avaient fait des enquêtes sur un groupe d'étudiants, dont la plupart étaient prêts à payer seulement deux ou trois dollars pour jouer.

Statistiquement, Joseph Bertrand avait déjà abordé le fait que la moyenne des gains dépendait fondamentalement du nombre de répétitions du jeu. William Feller démontrait formellement en 1937 une loi faible des grands nombres généralisée à certaines variables aléatoires sans espérance (i.e. « espérance infinie »). Cette loi conduit alors à proposer un enjeu

équitable par partie de $1 + \log_2(n)/2$, avec n , le nombre de parties convenu à l'avance (Voir D. Zajdenweber 1994).

Pour D'Alembert, l'espérance de gain de Pierre aurait été de :

$$E = 1/2 * 1 + 1/2 * 2/3 * 2 + 1/2 * 1/3 * 3/4 * 4 + 1/2 * 1/3 * 1/4 * 4/5 * 8 + \dots = 1/2 + 2/3 + 1/2 + 4/15 + \dots \approx 2,10.$$

D'Alembert dans son point 44 analysait ce que devrait être les contributions à l'espérance de gain de chaque gain possible. Il démontrait que la contribution maximale devait être pour le gain de 2 écus et que la décroissance des contributions suivantes devait être suffisamment forte pour que l'espérance du jeu soit finie. En effet, la contribution à l'espérance de gain d'un pile au premier coup est de 1/2, elle est de 2/3 pour le possible deuxième coup, puis à nouveau de 1/2, puis de 4/15... D'Alembert avait tout prévu, mais cela ne lui a pas suffi pour trouver cette loi, cette loi qui lui aurait permis de justifier un pari rationnel au jeu de Saint Pétersbourg d'environ deux écus.¹⁰

Bibliographie

- [1] Jean le Rond D'Alembert, *Croix ou pile*, Encyclopédie Méthodique, 1754.
- [2] Jean le Rond D'Alembert, *Opuscules Mathématiques*, tome 2, 1761.
- [3] Jean le Rond D'Alembert, *Doutes et questions sur le calcul des probabilités*, 1767.
- [4] Jean le Rond D'Alembert, *Sur le calcul des probabilités*, Mémoire 52, 1780.
- [5] Nicolas de Béguelin, *sur l'usage du principe de raison suffisante dans le calcul des probabilités*, 1768.
- [6] Joseph Bertrand, *Calcul des probabilités*, Gauthier-Villars, 1889.
- [7] Pierre Crépel, *D'Alembert, mathématicien des lumières*, Les Génies de la science n° 39, 2009.
- [8] Léo Gerville-Réache, *La loi du «croix ou pile» de D'Alembert*, Prix Tangente 2016, à paraître dans Tangente n°175, 2017.
- [9] Léo Gerville-Réache, *L'enjeu rationnel du paradoxe de St Pétersbourg*. hal-01399519, 1p. 2016.
- [10] Pierre-Simon Laplace, *Essai philosophique sur les probabilités*, 1774.
- [11] Haim Levy & Marshall Sarnat, *Portfolio and Investment Selection: Theory and Practice*, Prentice-Hall, 1984.
- [12] Bernard Parzysz, *Expérience aléatoire et simulation : le jeu croix ou pile*, IREM. n° 66, 2007.
- [13] Michel Paty, *D'Alembert et les probabilités*, Blanchard, Paris, p. 203-265, 1988.
- [14] Walter Penney, *Problem 95: Penney-Ante*, Journal of Recreational Mathematics, 1969.
- [15] André Totahasina, *L'introduction du concept de probabilité conditionnelle : avantages et inconvénients de l'arborescence*, Repères-IREM n°15, p. 93-117, 1994.
- [16] Daniel Zajdenweber, *Équité et jeu de Saint-Pétersbourg*, Revue économique n°45(1), p. 21-46, 1994.

⁷ Et cela, deux siècle avant Maurice Allais (à suivre...)
⁸ La modélisation probabiliste est l'art du maniement conjoint et rigoureux des principes de raison suffisante et insuffisante.
⁹ Précisément, Béguelin travaille avec des gains deux fois moindres que ceux de l'énoncé de D'Alembert et son calcul donnait 2,45 écus.

¹⁰ Personnellement, si l'on accepte l'hypothèse du jeu qui suppose que Paul dispose d'une fortune infinie, il doit en être de même pour Pierre. Ce dernier acceptera alors sans état d'âme de payer un enjeu infini puisque qu'à l'issue du jeu, quel qu'il soit, chacun repartira avec sa fortune de départ. C'est ça l'infini dénombrable du jeu de Saint Pétersbourg (voir Gerville-Réache 2016).