que sais-je ?

LES PROBABILITÉS

ALBERT JACQUARD

le point des connaissances actuelles

SES UNIVERSITAIRES DE FRANCE

Les probabilités

Les probabilités

ALBERT JACQUARD
Ancien élève de l'Ecole polytechnique
Docteur en Biologie humaine

Quatrième édition

32e mille

ISBN 2 13 039237 7

Dépôt légal — 1re édition : 1974
4e édition : 1986, février

108, boulevard Saint-Germain, 75006 Paris

Chapitre Premier

LE CONCEPT DE PROBABILITÉ

I. — Pascal et le problème des « partis » [1]

Le 29 juillet 1654, Pascal, malade, obligé de garder le lit, n'attend pas d'être rétabli pour écrire une longue lettre à Fermat ; il ne peut contenir son impatience à s'exprimer, à manifester sa joie ; son enthousiasme est tel qu'il ne parvient plus à maîtriser ses mots et doit terminer son exposé en latin « car le français n'y vaut rien ».

Il venait d'imaginer une méthode simple pour apporter une réponse à un problème posé par le chevalier de Méré : « Deux joueurs (disons Primus et Secondus) engagent chacun 32 pistoles dans un jeu de pile ou face ; empochera les 64 pistoles celui d'entre eux qui, le premier, aura obtenu 3 succès, consécutifs ou non. Ils jouent une première manche, Primus gagne ; ils sont à ce moment obligés de se séparer, la partie ne sera jamais terminée. Comment partager *équitablement* l'enjeu entre eux ? »

La partie n'a pas eu lieu ; on pourrait estimer que le contrat entre les deux joueurs est caduc ; ils doivent reprendre chacun leur enjeu.

[1] C'est-à-dire du « partage ».

Ce raisonnement est parfaitement justifié ; mais il ne tient pas compte d'un fait important : Primus a gagné une manche, son espoir de remporter la partie est plus grand que l'espoir que peut en avoir Secondus. Il peut paraître plus conforme à l'équité de répartir l'enjeu d'une façon qui corresponde aux « espoirs » de chacun. Mais comment faire ?

Fermat, dans une lettre parvenue à Pascal le 28 juillet, propose une solution fondée sur le calcul des combinaisons, mais, selon Pascal, « la peine est excessive » et il lui propose dès le lendemain une méthode tout autre, que voici :

LETTRE DE PASCAL A FERMAT

Le 29 juillet 1654.

MONSIEUR,

L'impatience me prend aussi bien qu'à vous et, quoique je sois encore au lit je ne puis m'empêcher de vous dire que je reçus hier au soir, de la part de M. de Carcavi, votre lettre sur les partis, que j'admire si fort que je ne puis vous le dire. Je n'ai pas le loisir de m'étendre, mais, en un mot, vous avez trouvé les deux partis des dés et des parties dans la parfaite justesse ; j'en suis tout satisfait, car je ne doute plus maintenant que je ne sois dans la vérité, après la rencontre admirable où je me trouve avec vous. [...]

Votre méthode est très sûre et est celle qui m'est la première venue à la pensée dans cette recherche ; mais parce que la peine des combinaisons est excessive, j'en ai trouvé un abrégé et proprement une autre méthode bien plus courte et plus nette, que je voudrais pouvoir vous dire ici en peu de mots : car je voudrais désormais vous ouvrir mon cœur, s'il se pouvait, tant j'ai de joie de voir notre rencontre. Je vois bien que la vérité est la même à Toulouse et à Paris.

Voici à peu près comme je fais pour savoir la valeur de chacune des parties, quand deux joueurs jouent, par exemple, en trois parties, et chacun a mis 32 pistoles au jeu :

Posons que le premier en ait deux et l'autre une ; ils jouent maintenant une partie, dont le sort est tel que, si le premier la gagne, il gagne tout l'argent qui est en jeu, savoir, 64 pistoles ; si l'autre la gagne, ils sont deux parties à deux parties, et par conséquent, s'ils veulent se séparer, il faut qu'ils retirent chacun leur mise, savoir, chacun 32 pistoles.

Considérez donc, Monsieur, que si le premier gagne, il lui appartient 64 ; s'il perd, il lui appartient 32. Donc s'ils veulent ne point hasarder cette partie et se séparer sans la jouer, le premier doit dire : « Je suis sûr d'avoir 32 pistoles, car la perte même me les donne ; mais pour les 32 autres, peut-être je les aurai, peut-être vous les aurez ; le hasard est égal ; partageons donc ces 32 pistoles par la moitié et me donnez, outre cela, mes 32 qui me sont sûres. » Il aura donc 48 pistoles et l'autre 16.

Posons maintenant que le premier ait deux parties et l'autre point, et ils commencent à jouer une partie. Le sort de cette partie est tel que, si le premier la gagne, il tire tout l'argent, 64 pistoles ; si l'autre la gagne, les voilà revenus au cas précédent, auquel le premier aura deux parties et l'autre une.

Or, nous avons déjà montré qu'en ce cas il appartient à celui qui a les deux parties, 48 pistoles : donc, s'ils veulent ne point jouer cette partie, il doit dire ainsi : « Si je la gagne, je gagnerai tout, qui est 64 ; si je la perds, il m'appartiendra légitimement 48 : donc donnez-moi les 48 qui me sont certaines au cas même que je perde, et partageons les 16 autres par la moitié, puisqu'il y a autant de hasard que vous gagniez comme moi. » Ainsi il aura 48 et 8, qui sont 56 pistoles.

Posons enfin que le premier n'ait qu'une partie et l'autre point. Vous voyez, Monsieur, que, s'ils commencent une partie nouvelle, le sort en est tel que, si le premier la gagne, il aura deux parties à point, et partant, par le cas précédent, il lui appartient 56 ; s'il la perd, ils sont partie à partie : donc il lui appartient 32 pistoles. Donc il doit dire : « Si vous voulez ne la pas jouer, donnez-moi 32 pistoles qui me sont sûres, et partageons le reste de 56 par la moitié. De 56, ôtez 32, reste 24 ; partagez donc 24 par la moitié, prenez-en 12 et moi 12, qui avec 32, font 44. » [...]

La lecture de ce texte peut sembler malaisée ; dessinons sous forme d'un « arbre » les divers cours qu'aurait pu suivre la partie après la victoire initiale de Primus. Si Primus gagne la seconde manche puis la troisième, la partie s'arrête et il empoche l'enjeu ; s'il gagne la seconde mais perd la troisième, il faut jouer une quatrième et éventuellement une cinquième manche ; après la cinquième manche l'un des joueurs en a obligatoirement gagné trois et la partie est jouée.

L'arbre des éventualités a finalement la forme indiquée par la figure 1 où les manches sont représentées par des traits surmontés d'un P ou d'un S selon le joueur gagnant :

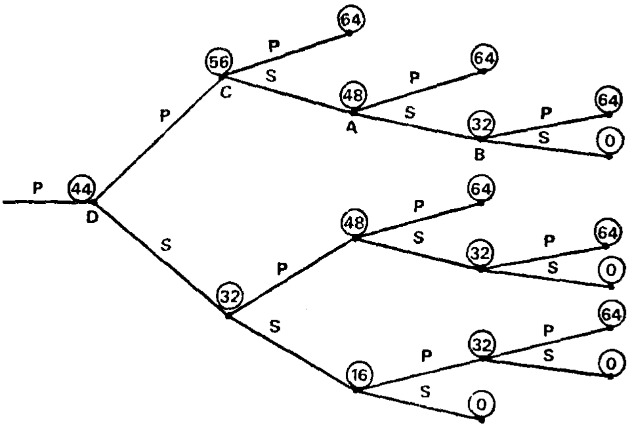

Fig. 1

On voit qu'au total 10 successions de manches étaient possibles, 6 d'entre elles aboutissent à la victoire de Primus, 4 à celle de Secondus (1).

La méthode proposée par Pascal consiste à explorer cet arbre *en commençant par la droite*. Dans le paragraphe « Posons que le premier en ait deux et

(1) Résistons à la tentation d'en conclure : « L'enjeu doit être partagé selon les proportions 6/10 et 4/10 ! » Le lecteur qui aborde le calcul des probabilités aura grand intérêt à interrompre ici quelques instants sa lecture et à découvrir pourquoi une telle conclusion serait incohérente avec ce que nous savons du jeu.

l'autre une ; ... », il se place dans une hypothèse qui correspond par exemple au point A de la figure 1. Considérons la situation en ce point ; Primus a gagné deux manches, Secondus une : deux cas sont possibles pour la manche suivante :

— ou Primus gagne et empoche l'enjeu de 64 pistoles ;
— ou Secondus gagne et la partie arrive au point B ;

en ce point B, les deux joueurs sont à égalité (deux manches partout) et leurs espoirs de remporter la partie sont identiques. L'espoir qu'a Primus de gagner correspond donc à 64 pistoles dans le premier cas, 32 dans le second ; comme ces deux éventualités sont « également probables », cet espoir correspond à la valeur $\frac{64+32}{2} = 48$ pistoles.

Dans les paragraphes suivants « le premier a deux parties et l'autre point : ... » et « le premier a une partie et l'autre point... », Pascal se place dans les hypothèses correspondant aux points C puis D de la figure et tient chaque fois un raisonnement semblable. Il trouve ainsi que l'espoir de gain de Primus en D est de 44 pistoles, ce qui résout le problème posé.

Sur la figure 1, nous avons indiqué dans un cercle, pour chacune des situations possibles, l'espoir de gain de Primus tel qu'on peut le calculer de proche en proche en se déplaçant sur le graphique de la droite vers la gauche.

II. — Remarques sur la méthode de Pascal

Il faut tout d'abord insister sur un point important. L'objet du discours dans le problème posé, et maintenant résolu sans « peine excessive », est la fin d'une partie définitivement interrompue ; non seulement cet objet est inconnu, mais il est inconnaissable ; bien plus, on sait fort bien qu'il n'existe pas, et n'existera jamais. Nous avons cependant pu dérouler un raisonnement rigoureux, capable d'emporter l'adhésion.

Cela a été possible grâce à l'introduction de deux concepts, celui d'*espérance* et celui de *probabilité*.

1. L' « espérance ». — Pour chacune des situations qui peuvent se produire au cours de la partie, nous avons estimé la valeur de l'espoir qu'avait chaque joueur de gagner. Autrement dit, nous avons admis que le joueur, dans la situation A par exemple, avait, face à un avenir incertain, un espoir de gain équivalent à celui qu'il aurait face à une loterie lui attribuant, à coup sûr, 48 pistoles.

Remarquons d'ailleurs que l'issue de la partie ne peut être qu'un gain de 64 pistoles ou d'aucune, jamais de 48. Il serait tout à fait erroné d'énoncer que « le joueur, lorsqu'il se trouve dans la situation A, a l'espoir de gagner 48 pistoles ». Le langage, qui a adopté le terme « espérance de la valeur » (en anglais *expected value*), traduit donc mal le concept évoqué, auquel correspondrait beaucoup mieux le terme « valeur de l'espérance ».

Le gain lui-même est un nombre qui, selon le résultat des « pile ou face » successifs, est égal soit à 0 soit à 64 pistoles ; il est variable ; il est aléatoire. Aux grandeurs, telles que ce gain, qui peuvent prendre diverses valeurs selon le résultat d'une épreuve où intervient le hasard, on donne le nom de *variables aléatoires*.

Remarquons enfin que l' « espérance » n'est pas une variable aléatoire ; nous avons pu la calculer avec précision selon la méthode proposée par Pascal sans avoir à connaître le résultat des manches successives. L'espérance est un nombre certain.

2. **Les probabilités.** — Pour évaluer cette espérance, on a considéré dans chaque cas les deux éventualités possibles (pile ou face) et l'on a fait la moyenne arithmétique des espérances attachées à ces éventualités ; cette opération était intuitivement justifiée par le fait que les deux issues possibles, pour chaque manche, ont « la même probabilité », ou comme le dit Pascal que « le hasard est égal ».

Que recouvre ce mot *probabilité* ? Il correspond à notre degré de confiance dans la réalisation d'un événement.

Si, comme dans le jeu de pile ou face, deux résultats seulement sont possibles et si je pense que la réalisation de l'un est aussi probable que celle de l'autre, je dis « qu'ils sont équiprobables », ou que « dans une moitié des cas l'un se produira, dans une moitié l'autre », ou, enfin, « qu'ils ont tous deux la probabilité 1/2 ».

Si l'un des événements est pratiquement certain et l'autre pratiquement impossible, je dis que « le premier se produira dans la totalité des cas, l'autre jamais », ou que « l'un a une probabilité proche de l'unité, l'autre une probabilité presque nulle ».

Il est clair que c'est autour de ce concept de probabilité que tout le raisonnement va s'articuler. Il faudra donc, au-delà d'une présentation qui jusqu'ici n'a fait appel qu'à l'intuition, rechercher dans ce domaine des définitions rigoureuses.

III. — Définitions formelles

Après cette approche purement intuitive, il importe de donner aux divers mots, que nous avons déjà employés de façon abusive, une signification précise qui nous permettra de raisonner avec rigueur.

1. Epreuve, résultats, événements. — Dans tout raisonnement probabiliste, nous considérons un *événement* E dont la réalisation dépend du *résultat* d'une *épreuve* qui n'a pas encore eu lieu ou dont l'issue n'est pas encore connue. Dans le cas du problème de Pascal, l'épreuve était, à chaque manche, le jet d'une pièce de monnaie ; le *résultat* de cette épreuve était le fait que la pièce montrait le côté « pile » ou « face », *l'événement* était le gain de la manche par Primus ou par Secondus.

Dans ce cas la distinction entre « résultat » et « événement » semble superflue car la correspondance entre eux est immédiate : Primus gagne si le résultat est « pile », Secondus si le résultat est « face ». Mais des exemples plus complexes peuvent être évoqués.

a) Tirons une carte au hasard dans un jeu de bridge ; cette *épreuve* a 52 résultats possibles. Le fait que la carte soit un carreau est un *événement* qui correspond à 13 de ces *résultats* ; le fait qu'elle soit un roi un *événement* qui correspond à 4 de ces *résultats*, etc.

b) Etudions la descendance d'un couple constitué de deux individus hétérozygotes pour un certain locus, c'est-à-dire possédant sur la partie de chromosome gouvernant un certain caractère deux gènes ayant des actions différentes, l'un A, l'autre a ; supposons de plus que l'action de A l'emporte sur celle de a (on dit alors que A est dominant et a récessif). Leur « génotype » à tous deux est Aa, mais en raison de la récessivité du gène a, son effet est masqué par l'effet du gène A ; les deux membres du couple présentent donc le caractère A : leur « phénotype » est [A].

L'*épreuve* consiste en l'émission par le père d'un spermatozoïde portant soit le gène A, soit le gène a, par la mère d'un ovule portant soit le gène A, soit le gène a. Quatre *résultats* sont possibles : r_1 (spermatozoïde A, ovule A) — r_2 (Sp. A, Ov. a) — r_3 (Sp. a, Ov. A) — r_4 (Sp. a, Ov. a).

L'*événement* étudié est, par exemple, le génotype de l'enfant ; il peut avoir trois formes : (AA) correspondant à r_1, (Aa) correspondant à r_2 et r_3, (aa) correspondant à r_4.

L'*événement* peut aussi être le phénotype de l'enfant, ce phénotype peut avoir deux formes :

[A] correspondant aux résultats r_1, r_2, r_3;
[a] correspondant au résultat r_4.

Finalement :

Nous désignons par épreuve *une expérience ou une observation réalisée dans des conditions bien définies, et dont le* résultat *est l'un des éléments d'un ensemble déterminé. Une épreuve est donc définie si l'on est en mesure de préciser :*

— *les conditions de l'expérience ou de l'observation ;*
— *l'ensemble* R *des résultats possibles* r_i.

Un événement E_1 *est lié à une épreuve si sa réalisation dépend du résultat de cette épreuve ; c'est-à-dire si l'on peut définir le sous-ensemble* R_1 *des résultats* r_i *qui entraînent la réalisation de cet événement.*

Cas particuliers. — Un événement est dit *élémentaire* s'il correspond à un et à un seul résultat de l'épreuve. (Ainsi, le gain d'une manche par Primus était un événement correspondant à un seul résultat de l'épreuve, le résultat « pile ».)

— Deux événements sont dits *incompatibles* lorsque leur réalisation est subordonnée à des résultats non communs de l'épreuve. Ainsi, dans le cas du choix d'une carte dans un jeu, les événements « la carte est un carreau » et « la carte est un trèfle » sont incompatibles ; par contre, les événements « la carte est un roi » et « la carte est un pique » ne sont pas incompatibles (puisque le résultat « roi de pique » leur est commun).

— Les événements $E_1, \ldots, E_n$ forment un *système complet* si tout résultat de l'épreuve entraîne la réalisation d'un et d'un seul de ces événements : ils sont incompatibles deux à deux et l'un d'entre eux est, nécessairement, réalisé.

Ainsi, les trois événements « la carte est un carreau », « la carte est un cœur », « la carte est de couleur noire » forment un système complet.

— Deux événements E_1 et E_2 sont dits *contraires* s'ils forment un système complet ; autrement dit, s'ils sont incompatibles et si l'un d'eux est nécessairement réalisé. On écrit alors : $E_2 = \overline{E}_1$.

Ainsi les événements « la carte est rouge », « la carte est noire » sont contraires ; il est équivalent d'énoncer « la carte est rouge » ou « la carte n'est pas noire ».

2. Probabilités. — Une épreuve étant définie, nous avons certaines informations ou certaines opinions sur son déroulement ; notre confiance en la réalisation de chacun des résultats possibles est plus ou moins élevée. Dans le cas du problème de Pascal, l'on a admis que chacun des deux résultats pile ou face a la même chance de se produire ou que, selon l'expression de Pascal, le « hasard est égal ».

De même, lorsque nous tirons une carte dans un jeu de bridge, nous pouvons admettre que ce jeu a été battu correctement, que chacune des cartes a une même chance d'être tirée ; lorsque, lors de la réalisation d'un spermatozoïde ou d'un ovule, l'un des deux gènes que possède l'individu est choisi pour être transmis à son descendant, nous admettons que chacun des gènes a la même chance d'être retenu.

Mais cette équiprobabilité des divers résultats n'est pas de règle ; nous pouvons posséder des informations qui nous amènent, au contraire, à estimer que certains résultats ont une chance meilleure que d'autres.

Finalement, nous disons que :

Une épreuve *est* probabilisable *si, compte tenu des opinions ou des informations que nous avons sur les conditions de son déroulement, nous sommes en mesure d'attacher à chacun de ses résultats possibles r_i un nombre p_i qui traduit le niveau de notre confiance dans sa réalisation.*

Par convention, nous choisissons des nombres p_i positifs d'autant plus grands que notre confiance dans le résultat r_i est plus élevée et tels que leur somme soit égale à 1 :

$$p_i \geqslant 0 \qquad \sum_{i \in \mathrm{R}} p_i = 1 \quad (^1). \tag{1}$$

Les nombres p_i sont les *probabilités* des résultats r_i.

La probabilité $\mathrm{P}(\mathrm{E}_1)$ d'un événement E_1 correspondant aux résultats r_k, r_l et r_m de l'épreuve est, par définition :

$$\mathrm{P}(\mathrm{E}_1) = p_k + p_l + p_m$$

(1) Rappelons que le symbole $\sum\limits_{i \in \mathrm{R}} p_i$ signifie : somme des nombres p_i correspondant aux indices i qui appartiennent à l'ensemble R (ici l'ensemble des résultats possibles de l'épreuve).

ou de façon générale :

$$P(E_1) = \sum_{i \in R_1} p_i \tag{2}$$

où R_1 est le sous-ensemble de résultats entraînant la réalisation de l'événement E_1.

Cas particuliers. — La probabilité d'un événement *élémentaire* est égale à la probabilité du résultat correspondant.

— Un événement E est *certain* s'il est réalisé quel que soit le résultat de l'épreuve ; autrement dit, le sous-ensemble R_c des résultats conditionnant E est confondu avec l'ensemble R des résultats possibles. Compte tenu des relations (2) et (1), on peut écrire :

$$P(E_{\text{certain}}) = 1.$$

— Un événement E_i est *impossible* si aucun des résultats de l'épreuve n'entraîne sa réalisation. L'ensemble R_i est donc l'ensemble vide, par conséquent :

$$P(E_{\text{impossible}}) = 0.$$

Exemples. — *a)* Supposons que nous ayons battu les cartes d'un jeu de bridge de façon telle que chacune des cartes ait la même chance d'être tirée. Les 52 résultats possibles sont équiprobables. On a donc :

$$p_1 = p_2 = \ldots = p_{52} = \frac{1}{52}.$$

L'événement E_C : « la carte est un carreau » est réalisé pour 13 résultats de l'épreuve, sa probabilité est donc :

$$P(E_C) = 13 \times \frac{1}{52} = \frac{1}{4}.$$