MATHÉMATIQUES ET SCIENCES HUMAINES

M. BARBUT

De Pascal à Savage. Un chapitre de l'algèbre linéaire : le calcul des probabilités (cas fini)

Mathématiques et sciences humaines, tome 15 (1966), p. 15-28

<http://www.numdam.org/item?id=MSH_1966__15__15_0>

M. BARBUT

DE PASCAL A SAVAGE. UN CHAPITRE DE L'ALGEBRE LINEAIRE : LE CALCUL DES PROBABILITES (cas fini)*

L'exposé qui suit n'a d'autre prétention que de suggérer aux enseignants une voie moins classique (bien que très ancienne, 1654, et assez naturelle) que celle qui est suivie dans la plupart des manuels pour l'initiation au Calcul des Probabilités. (Une première version (5), écrite il y a quelques années, est maintenant épuisée).

Ce mode d'exposition a, à mon avis, quelques avantages majeurs: d'abord, celui d'insérer le Calcul des Probabilités dans le contexte général de l'algèbre linéaire, partout présente, et qui fait l'essentiel de tous les programmes de mathématiques; ensuite celui d'atténuer les difficultés que l'on rencontre auprès des étudiants lors de l'introduction de la notion de variable aléatoire, quand on a commencé par définir d'abord les probabilités. Ici, la variable "incertaine", puis la variable aléatoire sont les notions dont on part d'entrée de jeu; et la notion de base celle d'espérance, c'est-à-dire de forme linéaire positive sur un vectoriel de fonctions numériques: l'intégrale.

Enfin, cette méthode permet d'aborder le Calcul des Probabilités puis surtout la Statistique Mathématique comme ce qu'ils sont en fait comme domaine des mathématiques appliquées: un chapitre de la théorie mathématique des décisions.

I - Le vectoriel des Paris

a) Un pari, c'est la donnée d'un ensemble A d'éventualités, et, pour chaque éventualité $i \in A$, d'un nombre x_i : le gain (ou la perte), l'évaluation des conséquences, pour le parieur, de l'éventualité i.

Par exemple, un pari à pile ou face. $A = \{1, 2\}$, $x_1 = 3$, $x_2 = -5$: le parieur gagne 3 francs si c'est pile, perd 5 francs si c'est face.**

* Conférence faite au premier stage international du Centre Belge de Pédagogie de la Mathématique, Knokke, le 21 Mai 1966.

**Attention, actuellement, et jusqu'à ce que cette notion soit introduite, nous ne savons pas ce qu'est la probabilité de pile, ou celle de face.

Chaque pari x apparait ainsi comme une liste de nombres, autant que d'éléments de A si A est fini, ce que nous supposerons jusqu'à nouvel ordre :

$$x = (x_1, x_2, \ldots, x_n)$$

Si l'on joue aux dés, avec deux dés, $n = 36$;
si l'on joue au tiercé, et qu'il y ait dix chevaux partants, $n = 10 \times 9 \times 8 = 720$, etc...

Mais il n'y a pas que dans les jeux de hasard que l'on rencontre des paris: souscrire un contrat d'assurance, par exemple, peut être assimilé à tenir un pari dans lequel A est l'ensemble des risques couverts, le nombre x_i la conséquence financière du sinistre i : somme remboursée par l'assureur moins la perte afférente à ce sinistre.

Plus généralement on est dans une situation de pari, chaque fois que l'on doit prendre une décision dont les conséquences, évaluées par des nombres, dépendent d'éventualités dues au hasard et sur lesquelles nous sommes sans aucun pouvoir.

Les nombres x_i par lesquels sont évaluées les conséquences du pari, ne sont pas nécessairement les valeurs nominales (en francs, ou dans toute autre unité monétaire) de la perte ou du gain encouru; ce sont les <u>utilités</u> (au sens des économistes) de ces conséquences pour le parieur.

b) Le principal des problèmes qui se posent à propos des paris est un problème de comparaison et de choix. Si l'on m'offre de tenir un pari tel que le pari à pile ou face donné plus haut en exemple, j'ai le choix entre tenir, ou ne pas jouer. Ne pas jouer, c'est encore un pari, le pari nul: quoiqu'il arrive, le résultat en est nul. Et il me faut comparer ces deux paris pour arriver à une décision rationnelle. Or, cette comparaison n'est pas facile: si c'est l'éventualité 1 qui se produit, c'est tenir le pari qui est l'acte le plus avantageux (je gagne 3 francs); mais si c'est l'éventualité 2, c'est de ne rien faire (j'évite une perte de 5 francs). Or, je suis dans l'incertitude complète quant à ce qui va arriver, et c'est maintenant qu'il me faut choisir.

De même, si l'on m'offre pour un même ensemble A d'éventualités de souscrire à l'un ou l'autre des deux contrats d'assurance :

$$x = (x_1, x_2, \ldots, x_n)$$

$$y = (y_1, y_2, \ldots, y_n)$$

il arrivera en général que pour certaines éventualités x soit plus avantageux que y ($x_i > y_i$), et que pour d'autres ce soit le contraire ($x_i \leqslant y_i$).

Celui qui "calcule", comme on dit, sa décision, pourra, dans un cas comme celui-ci, examiner en quoi diffèrent les deux contrats; nous avons supposé que les nombres x_i, par lesquels sont évaluées les conséquences sont des utilités: nous supposerons en outre que l'on peut d'une part les additionner entre eux (ou les soustraire), et les soumettre à des changements d'unité (des homothéties); en pratique, on les considère toujours comme appartenant au corps des réels ou à celui des rationnels.

Examiner la différence entre x et y prend alors le sens précis de calculer la liste :

$$x - y = (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)$$

On peut également lorsqu'on calcule sa décision, envisager de souscrire à la fois aux deux contrats, ce qui correspond à l'opération :

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

L'ensemble de tous les paris possibles sur un même ensemble A d'éventualités est ainsi muni d'une addition, avec les propriétés connues de l'addition vectorielle; en particulier, le pari nul, c'est-à-dire le vecteur nul :

$$0 = (0,0,\ldots,0)$$

y joue, on l'a vu, un rôle important, puisque c'est à lui que l'on devra en définitive comparer, lors d'une décision, tout autre pari.

Les homothéties ont également un sens dans ce type de calcul; le pari :

$$\lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

homothétique de x dans le rapport λ est tout simplement celui qui se déduit de x par changement d'unité de rapport λ sur les utilités. Par exemple, si les utilités x_i sont exprimées en francs français, et qu'on veuille les exprimer en francs belges, $\lambda = 9{,}8475$, le 23 Mai 1966.

En résumé, l'ensemble des paris possibles sur un même ensemble A d'éventualités est le vectoriel de dimension égale au nombre d'éléments de A, ayant pour corps de scalaires le corps des réels (ou des rationnels).

L'addition $x + y$ de deux vecteurs a pour signification : tenir à la fois les deux paris représentés par les vecteurs x et y respectivement. L'homothétie: $x \longrightarrow \lambda x$ de rapport λ correspond au changement d'unités sur les utilités.

Ce vectoriel est muni de sa structure d'ordre partiel :

$$x \geqslant y \Longleftrightarrow \forall_i \in A, \quad x_i \geqslant y_i$$

cette inégalité signifie qu'en toute éventualité le pari représenté par le vecteur x est au moins aussi avantageux que celui qui est représenté par y.

II - Une forme linéaire positive : l'espérance.

a) Le problème fondamental, avons-nous dit, c'est celui de la comparaison et du choix entre paris, ou entre paris et valeurs certaines. C'est plus précisément celui de la valeur qu'il convient d'attacher à un pari, son utilité (toujours au sens des économistes): les paris sont en effet des "biens" qui se vendent et s'achètent couramment. Lorsque j'achète un billet de loterie, lorsque je souscris à un contrat d'assurance, c'est bien un pari que j'achète.

Reprenons un exemple très simple, celui du pari $x = (3, -5)$ à deux éventualités; si l'on me demande de payer deux francs le droit de parier, il est probable que je refuserai; si l'on me donne un franc si je tiens le pari, peut-être accepterai-je: ce qui signifie que pour moi la valeur, l'utilité de ce pari se situe entre -1 et 2.

Postulons que l'on puisse effectivement fixer la valeur, l'utilité d'un pari: ce sera la somme pour laquelle il m'est indifférent, soit de tenir le pari, soit d'avoir certainement cette somme. Nous appelons espérance d'un pari x, notée $E(x)$, son utilité, cette valeur d'équilibre entre pari et certitude.

Cette interprétation de l'espérance d'un pari comme son utilité, sa valeur, n'est pas nouvelle; c'est par exemple celle de von Neumann dans (1); nous verrons plus loin qu'elle est également présente chez Pascal. Et c'est une valeur de l'espérance qu'exprime le dicton: "Un tiens vaut mieux que deux tu l'auras".

b) L'espérance $E(x)$ d'un pari x, disions-nous; mais, bien entendu, il faut pouvoir calculer l'espérance de n'importe quel pari, tout au moins de tous ceux qui admettent le même ensemble A d'éventualités, de sorte que la comparaison entre deux paris x et y se réduise à la comparaison, toujours possible, entre les deux nombres $E(x)$ et $E(y)$.

L'espérance doit donc être un opérateur appliquant le vectoriel des paris dans le corps représentant l'échelle des utilités. Si celui-ci est le corps $\mathbb{R}$ des réels, et si nous notons $\mathbb{P}_n$ le vectoriel de dimension n des paris sur un ensemble A de n éventualités, l'espérance E est une application de $\mathbb{P}_n$ dans $\mathbb{R}$:

$$E : \mathbb{P}_n \longrightarrow \mathbb{R}$$

Il est bien évident, du point de vue algébrique, que cette application doit être linéaire et positive, c'est-à-dire satisfaire aux trois conditions:

1) $E(x + y) = E(x) + E(y)$

2) $E(\lambda x) = \lambda E(x)$

3) $x \geqslant 0 \Longrightarrow E(x) \geqslant 0$

3) équivaut d'ailleurs, en présence de 1) à ce que E soit monotone :

$$x \geqslant y \Longrightarrow E(x) \geqslant E(y)$$

Mais ces trois conditions s'imposent également dans notre interprétation de l'espérance d'un pari comme son utilité, comme son prix.

La positivité 3), exprime en effet que si un pari x a en toute éventualité une conséquence positive, sa valeur (son espérance) doit être positive. La condition 2) ne fait qu'assurer l'homogénéité des unités choisies pour évaluer les utilités: l'espérance doit s'exprimer dans la même unité que les conséquences attendues du pari; si je fais un changement d'unités de rapport λ sur celles-ci je dois faire le même changement d'unités sur l'espérance. Quant à la condition 1), elle dit que le prix à payer pour acheter à la fois deux paris (deux billets de loterie; deux contrats d'assurance) doit être la somme des prix à payer pour chacun d'entre eux. Ce n'est peut-être pas toujours vrai dans la pratique (il peut par exemple y avoir des "ristournes"), mais nous l'admettrons comme première approximation, quitte à discuter plus loin cette hypothèse.

On sait que 1), 2), et 3) ne sont pas indépendantes; mais notre but n'est pas de donner une axiomatique minimale, c'est de faire comprendre le mécanisme de l'espérance.

Aux trois conditions écrites, il faut en ajouter une quatrième qui assure que l'espérance d'un pari coïncide bien avec l'utilité de celui-ci dans le cas où cette utilité est la même en toute éventualité: les cas de certitude où, quoiqu'il arrive, les conséquences sont les mêmes. Nous poserons donc :

4) Si $x_1 = x_2 = \dots = x_n = a$, alors
$E(x) = a$.

Nous venons de définir les conditions auxquelles doit satisfaire l'espérance; il reste à la construire.

III - Distributions de probabilité et variables aléatoires (ou aléas numériques).

a) Chaque pari x fait correspondre à chaque éventualité $i \in A$ un nombre réel x_i; c'est une application de A dans $\mathbb{R}$.

Certaines des valeurs x_i peuvent être égales entre elles (c'est presque toujours le cas dans les loteries: à la loterie nationale, par exemple, l'ensemble A est très vaste, c'est celui de tous les numéros possibles; mais le nombre des lots est restreint, de l'ordre de la dizaine). Les valeurs prises par les x_i constituent un ensemble $(y_1, y_2, \dots, y_k)$ de nombres $(k \leqslant n)$, et A est partitionné en k classes $A_1, A_2, \dots, A_k$, où :

$$i \in A_j \Longleftrightarrow x_i = y_j$$

Désignons par φ_B, où B est une partie quelconque de A, la "fonction caractéristique" de B :

$$\varphi_B i = 1 \quad \text{si} \quad i \in B$$

$$\varphi_B i = 0 \qquad \text{sinon}$$

En termes de paris, φ_B est le pari dans lequel je gagne 1 franc si se produit l'une des éventualités de B, et rien sinon.

On a alors l'égalité (vectorielle) :

$$x = y_1\varphi_{A_1} + y_2\varphi_{A_2} + \dots + y_k\varphi_{A_k}$$

Où $y_1, y_2, \dots y_k$ sont des scalaires, et $\varphi_{A_1}, \varphi_{A_2}, \dots, \varphi_{A_k}$ des vecteurs (des paris). En effet, chaque $i \in A$ appartient à une seule des classes; s'il appartient à la classe A_j, on a :

$$x_i = y_j \text{ et } \varphi_{A_r} i = 0 \text{ si } r \neq j, \ \varphi_{A_j} i = 1$$

E devant être linéaire, on doit avoir (conditions 1 et 2) :

$$E(x) = y_1\, E(\varphi_{A_1}) + y_2\, E(\varphi_{A_2}) + \dots + y_k\, E(\varphi_{A_k})$$

Mais pour chaque partie B de A, φ_B est entièrement définie par la donnée de B, et par conséquent $E(\varphi_B)$ est une fonction de B. Posons :

$$E(\varphi_B) = P(B)$$

Avec ce changement de notation :

$$E(x) = y_1 P(A_1) + y_2 P(A_2) + \dots + y_k P(A_k)$$

et l'espérance de x se calcule en faisant la somme des valeurs obtenues y_j pondérées par les coefficients $P(A_j)$. Il suffit de connaître, pour toute partie B de A, l'espérance P(B) du pari φ_B pour être en mesure de calculer l'espérance de n'importe quel pari.

b) Le nombre P(B), espérance du parieur dans le pari qui lui donne un franc (une unité de son échelle d'utilité) dans toutes les éventualités appartenant à B, et rien sinon, peut être considéré comme une mesure du "degré de probabilité" attaché par la parieur à ce que les probabilistes appellent l'évènement B. Nous dirons que par définition $P(B) = E(\varphi_B)$ est pour chaque partie B de A (pour chaque "évènement" B) la <u>probabilité de</u> B.

A chaque partie de A correspond sa probabilité; l'application P de l'ensemble des parties de A dans $\mathbb{R}$ n'est pas quelconque. Comme, quel que soit B, $\varphi_B \geqslant 0$ (φ_B ne prend que les valeurs 0 et 1), on a

1') $E(\varphi_B) = P(B) \geqslant 0$

D'autre part, φ_A est le pari qui en toute éventualité donne le résultat 1. D'après la quatrième condition sur les espérances :

2') $E(\varphi_A) = P(A) = 1.$

Enfin, si B et C sont deux parties disjointes de A, on a :

$$\varphi_{B\cup C} = \varphi_B + \varphi_C$$

d'où, d'après la linéarité de l'espérance :

$$E(\varphi_{B\cup C}) = P(B\cup C) = E(\varphi_B) + E(\varphi_C) = P(B) + P(C)$$

3') $B \cap C = \emptyset \Longrightarrow P(B\cup C) = P(B) + P(C)$

Nous retrouvons bien ainsi, pour la probabilité définie comme espérance de paris particuliers, les paris 0-1, les axiomes classiques de Kolmogorov (L'axiome des probabilités composées apparaîtra plus tard).

c) Appelons $p_1, p_2, \dots, p_n$ les probabilités des évènements $\{1\}, \{2\}, \dots, \{n\}$; ces nombres sont tous positifs ou nuls. D'après 3'), pour tout évènement B,

$$P(B) = \sum_{i \in B} p_i$$

et d'après 2')

$$\sum_{i=1}^{n} p_i = P(A) = 1$$

Toute suite $p = (p_1, p_2, \ldots, p_n)$ de nombres positifs (ou nuls) et de somme 1 constitue ce qu'on appelle une distribution de probabilité sur A. Au moyen de p, l'espérance $E(x)$ d'un pari $x = (x_1, x_2, \ldots, x_n)$ peut s'écrire :

$$E(x) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$$

Remarquons qu'on aurait pu obtenir directement cette expression de $E(x)$ en décomposant x sous la forme :

$$x = x_1 \varphi\{1\} + x_2 \varphi\{2\} + \ldots + x_n \{n\}$$

Les paris $\varphi\{i\}$ constituent une base du vectoriel des paris :

$$\varphi\{1\} = (1, 0, 0, \ldots, 0) \text{ etc....}$$

Mais le passage par les valeurs y_j prises par la fonction x de A dans $\mathbb{R}$ se prête mieux aux généralisations ultérieures aux cas où A est infini: introduction de l'intégrale de Lebesgue-Stieltjes*. De plus, cela correspond à un mode très usuel de calcul des moyennes.

d) Un pari x : une liste de nombres $(x_1, x_2, \ldots, x_n)$ une application

$$x : A \longrightarrow \mathbb{R}$$

de l'ensemble des éventualités dans le corps ordonné des réels. On dit que c'est une "variable incertaine" (il vaudrait mieux dire "fonction incertaine") si l'on veut indiquer la signification de cette fonction numérique: les valeurs x_i qu'elle prend sont les résultats, incertains, d'éventualités commandées par le hasard.

A partir du moment où l'on a une distribution de probabilité $p = (p_1, p_2, \ldots p_n)$ sur l'ensemble A d'éventualités l'incertitude change de caractère, elle est mesurée par p; le couple (x,p) est une variable aléatoire, ou mieux, un aléa numérique.

Pour un même ensemble A d'éventualités, il y a ainsi une infinité d'aléas numériques: leur ensemble est le produit cartésien $P_n \times S_n$, où P_n est le vectoriel des paris (le vectoriel de dimension n sur $\mathbb{R}$), et S_n le simplexe de dimension n.

IV - Espérances (et probabilités) conditionnelles: Pascal.

a) Jusqu'ici, nous avons considéré les paris sur un même ensemble d'éventualités A; ceci est insuffisant pour les besoins pratiques, car cet ensemble, du fait des informations qu'on recueille, de l'histoire qui se déroule, est, dans bien des cas, en perpétuelle évolution.

* Voir, par exemple (9).

Précisons, en nous plaçant dans le cas d'un jeu de hasard qui se déroule en plusieurs coups; à un certain moment, l'ensemble des éventualités est A, et mon pari est x, une application de A dans $\mathbb{R}$.

Si l'on joue un coup, l'ensemble A d'éventualités est partitionné en classes A_1, A_2, ..., A_r, selon le résultat de ce coup, de sorte qu'au coup suivant l'ensemble d'éventualités sera réduit soit à A_1, soit à A_2, ..., soit à A_r.

Par exemple, si l'on lance trois fois de suite un dé à six faces, A est, avant que l'on commence, l'ensemble des 216 triplets (u, v, w) de nombres entiers de 1 à 6. A l'issue du premier coup, on se trouvera avec pour nouvel ensemble d'éventualités une et une seule des six classes A_1, A_2, ..., A_6 : A_1 est l'ensemble des 36 triplets commençant par 1, A_2 celui des triplets commençant par 2, etc....

A l'issue d'un coup, l'ensemble d'éventualités se trouve donc réduit de A à l'une des classes, A_j, de la partition correspondante de A; quant au pari x (à la fonction x de A dans $\mathbb{R}$), il se trouve réduit à sa restriction x_{A_j} à A_j. Si l'on préfère le langage vectoriel, x_{A_j} est la projection de x sur le sous-espace engendré par A_j.

L'espérance de x_{A_j} (qui est une forme linéaire positive de l'espace des applications de A_j dans $\mathbb{R}$) est ce qu'on appelle l'<u>espérance conditionnelle</u> de x, si A_j, considéré comme évènement de A, se produit. Notons la $E_{A_j}(x_{A_j})$, où E_{A_j} est l'opérateur espérance pour l'ensemble A_j d'éventualités; notons de même E_A l'espérance lorsque A est l'ensemble d'éventualités.

C'est ici qu'intervient ce que nous appellerons le "principe de Pascal" : $E_{A_j}(x_{A_j})$ est la valeur du pari x_{A_j} ; ce qui résulte du coup que nous jouons, avec A comme ensemble d'éventualités, ce sont les paris x_{A_1}, x_{A_2}, ..., x_{A_r}, chacun ayant pour utilité, pour valeur (c'est le mot même employé par Pascal) son espérance. Par conséquent, nous pouvons écrire, en appliquant la formule générale du calcul de l'espérance :

$$E_A(x) = \sum_{j=1}^{r} P_A(A_j)\, E_{A_j}(x_{A_j})$$

Où $P_A(A_j) = E_A(\varphi_{A_j})$ est la probabilité lorsque A est l'ensemble d'éventualités de l'évènement A_j.

b) Autre formulation : φ_{A_j} désignant la fonction caractéristique dans A de A_j (φ_{A_j} vaut 1 pour les éléments de A_j, 0 pour les autres), nous avons :

$$x = x\,\varphi_{A_1} + x\,\varphi_{A_2} + \dots + x\,\varphi_{A_r}$$

Le produit $x\,\varphi_{A_j}$ coïncide avec x_{A_j} sur A_j, et vaut zéro ailleurs; l'égalité ci-dessus est une expression du vecteur x comme somme de ses projections sur les sous-espaces engendrés par $A_1, A_2, \ldots, A_r$ respectivement.

Si $j \neq k$, posons : $E_{A_k}(x \cdot \varphi_{A_j}) = 0$

($x\,\varphi_{A_j}$ est nulle sur A_k).

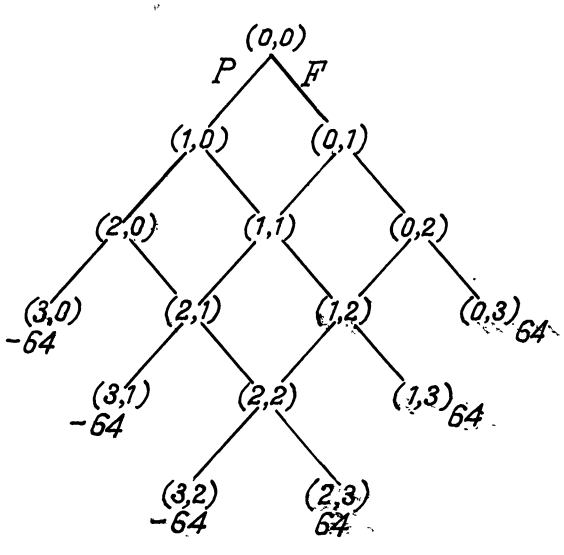
Le principe de Pascal, sa règle d'enchaînement des espérances, peut alors s'écrire :

$$E_A = \sum_{i=1}^{r} P_A(A_j)\, E_{A_j}$$

ce qui signifie que l'opérateur espérance, si A est l'ensemble d'éventualités, est une combinaison linéaire des opérateurs conditionnels E_{A_j}, avec pour coefficients les probabilités (si A) des évènements A_j.

c) J'ai dit "principe de Pascal". Lisons en effet sa lettre du 29 Juillet 1654 à Fermat, à propos du problème des partis.

Le problème, d'abord: deux joueurs jouent à pile ou face "en trois parties et chacun a mis 32 pistoles au jeu". Celui qui a parié sur pile empoche les 64 pistoles dès que pile sera apparu trois fois au cours d'une suite de coups, et inversement pour celui qui a parié sur face. On voit que cinq coups au plus suffisent à décider, et le jeu peut se schématiser par le réseau ci-contre, où l'on a indiqué pour chaque situation le "score" de chacun des joueurs et aux sommets terminaux le gain de celui qui a parié sur face.



"Voici à peu près, dit Pascal, comme je fais pour savoir la valeur de chacune des parties ....."

Soulignons le mot <u>valeur</u>, qui indique bien ce dont il s'agit: calculer l'espérance d'un pari dans le sens que nous avons donné ici à l'espérance, celle de l'utilité (selon le jargon moderne) d'un pari.

"Posons que le premier en ait deux et l'autre une;..." Ils sont dans la situation marquée par le sommet (1,2) du graphique, si le premier est le parieur de face. "Ils jouent maintenant une partie, dont le sort est tel que, si le premier la gagne, il gagne tout l'argent qui est au 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".

Nos joueurs sont dans la situation (1,2) qui peut les amener aux situations finales, (3,2), (2,3), (1,3). Pour eux, et dans nos notations :

$$A = \{(3,2),(2,3),(1,3)\}$$

est l'ensemble d'éventualités à ce coup; s'ils le jouent, A est partitionné en :

$$A_1 = \{(3,2),(2,3)\} \quad A_2 = \{(1,3)\}$$

selon que le hasard les conduit à la situation (2,2) ou à la situation A_2.

L'espérance (la valeur) si A_2 est réalisée, du parieur de face est la somme certaine 64; son espérance si A_1 est de 32, car, dira plus loin Pascal, "le hasard est égal".

Mais remarquons le petit membre de phrase : "..... s'ils veulent se séparer, il faut qu'ils retirent ....". Dans la situation (2,2), nos joueurs peuvent soit continuer à jouer, et remettre la décision du partage au hasard, soit ne pas jouer: la somme qu'ils doivent alors retirer est précisément leur espérance (estimée ici à 32 pistoles). Cette espérance est bien ici la valeur certaine qui équivaut pour eux à tenir le pari, le prix du pari.

Et la suite est encore plus nette :
"Considérez donc, Monsieur, que si le premier gagne, il lui appartient 64; s'il perd, il lui appartient 32". "Il lui appartient 32", son espérance si A_2 est ici encore la valeur pour lui de la situation A_2. "Donc, s'ils ne veulent 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 sures". "Il aura donc 48 pistoles et l'autre 16".

Ici, Pascal dit pratiquement tout en quelques lignes. D'abord, l'addition de deux paris sur le même ensemble $\{A_1, A_2\}$ d'éventualités: le pari (32, 32) et le pari (32, 0); et l'addition correspondante des espérances: l'espérance du pari global est la somme des deux espérances.

Mais pour la première, il s'agit d'un pari dans lequel j'ai 32 en toute éventualité (mais l'un de ces 32 est lui-même une espérance, non une somme certaine). Son espérance est 32: c'est la quatrième des conditions que nous avons imposées à l'espérance.

Et le résultat final s'exprime bien par

$$E_A = p(A_1)\, E_{A_1} + p(A_2)\, E_{A_2}$$

Ici, $p(A_1) = p(A_2) = \frac{1}{2}$, $E_{A_1}(x) = 32$, $E_{A_2}(x) = 64$.

Ayant calculé l'espérance (il dit le parti, c'est-à-dire le partage) en (1, 2), Pascal le calcule ensuite de proche en proche en (0, 2) et en (0,1) en appliquant toujours les mêmes principes: addition des paris, linéarité de l'espérance, formule des espérances conditionnelles, espérance des paris donnant le même résultat en toute éventualité.

d) De ce que nous avons appelé ici le principe (ou la règle) de Pascal, on déduit aisément le 4ème axiome classique du Calcul des Probabilités, celui des "probabilités composées". Nous avons :

$$E_A(x) = \sum P_A(A_j)\, E_{A_j}\, (x_{A_j})$$

Supposons que: $x = \varphi_B$, où B est une partie quelconque de A; alors la formule s'écrit:

$$P_A(B) = \sum P_A(A_j)\, P_{A_j}\, (B)$$

$P_{A_j}(B) = P_{A_j}(B \cap A_j)$ est la probabilité conditionnelle de B si A_j.

Particularisons encore: A est partitionné en deux classes seulement, X et sa partie complémentaire Y; et posons $B = X \cap Z$, où Z est une partie quelconque de A.

$$P_A(X \cap Z) = P_A(X)\, P_X(X \cap Z) + P_A(Y)\, P_Y(X \cap Z)$$

Mais

$$P_Y(X \cap Z) = P_Y(X \cap Z \cap Y) = P_Y(\emptyset) = 0$$

$$P_X(X \cap Z) = P_X(Z)$$

Et il reste l'égalité :

$$P_A(X \cap Z) = P_A(X)\, P_X(Z)$$

qui est celle des probabilités composées. Cette égalité peut encore s'écrire, si $P_A(X) \neq 0$,

$$\frac{P_A(X \cap Z)}{P_A\,(X)} = P_X(Z)$$

forme sous laquelle apparaît sa signification essentielle: la distribution de probabilité P_A, c'est, pour l'ensemgle A d'éventualités, une façon de traduire l'opinion du parieur sur la vraisemblance des évènements (parties) de A; l'égalité ci-dessus nous dit comment il doit modifier son opinion lorsque de A, l'ensemble d'éventualités,se réduit à un sous-ensemble X de probabilité non-nulle.

Remarque: Le troisième axiome de Kolmogorov, celui de l'additivité des probabilités d'évènements disjoints, étant lui aussi un cas particulier de la règle pascalienne d'enchaînement des espérances, une autre voie de présentation serait de partir d'emblée de paris dont les résultats sont eux-mêmes des paris, chacun évalué par son espérance, et toutes les espérances étant conditionnelles.

C'est à cela que nous invite Pascal; mais je pense que pour des débutants il est préférable d'opérer progressivement, comme je l'ai fait ici.

V - Quelques remarques en terminant.

a) ***Probabilités "subjectives" et rôle de la statistique.***

Pour avoir l'espérance, il suffit de se donner une distribution de probabilité

$$p = (p_1, p_2, \dots, p_n)$$

sur l'ensemble des éventualités; les p_i sont arbitraires, pourvu qu'ils satisfassent aux contraintes :

$$\forall i, \; p_i \geqslant 0 \; ; \quad \sum p_i = 1$$

Le parieur les estimera au mieux de son expérience, ce qu'il sait des conditions du pari, etc... En ce sens, les probabilités telles qu'elles apparaissent dans l'introduction que nous avons esquissée sont "subjectives"; deux parieurs pourront différer dans leurs estimations personnelles des p_i.

L'axiomatique des espérances nous assure de l'existence d'une mesure (et même d'une foule de mesures) des "degrés de vraisemblance" des évènements d'un même ensemble d'éventualités. Une chose est, dans une théorie, d'avoir une mesure; une autre en est de posséder des instruments de mesure "objectifs". Par exemple, la géométrie euclidienne assure que la pièce dans laquelle nous nous trouvons a une certaine hauteur de plafond; mais deux observateurs, s'ils ne disposent pas d'instrument tel qu'un mètre, pourront différer dans l'estimation de cette hauteur.

Le rôle de la statistique mathématique (plus particulièrement dans son chapitre "Estimation") c'est précisément de fournir une métrologie au Calcul des Probabilités, de façon que nos deux parieurs qui divergeaient au départ dans leurs estimations, verront, si tous deux acceptent

1 - de se soumettre à un certain type d'expérience (tirages dans une urne)

2 - les règles du Calcul des Probabilités

leur opinion converger vers une même estimation.

De là, on voit que dans l'expression "Calcul des Probabilités" c'est le mot calcul qui est important, plus que le mot probabilités.

b) ***L'hypothèse de linéarité n'est, comme toujours, qu'une approximation.***

Dès le début du 18ème siècle, les paradoxes du type "paradoxe de Saint Pétersbourg" ont mis en évidence que la règle d'évaluation proposée par Pascal pour les paris, fondée sur la linéarité, se heurte à de graves difficultés si on s'en sert sans précaution. Ces paradoxes sont ceux dans lesquels on considère un pari entre un riche, et un pauvre qui n'a par exemple, que 5 francs pour toute fortune; le pari à pile ou face dans lequel il perdrait 100 francs si pile, et il gagnerait 200 francs si face, a une espérance positive pour le pauvre (si le hasard est égal entre pile et face), et pourtant il est douteux que celui-ci accepte d'y risquer sa fortune.

Remplacer les valeurs en numéraire par des utilités ne résoud en général pas la difficulté. Car elle est insoluble.

Elle tient à ce que, comme tous les "modèles" linéaires, aussi bien ceux des sciences physiques que des sciences économiques ou sociales, le calcul des espérances, qui est le "modèle" d'une certaine classe de décisions en cas d'incertitude, n'est qu'une approximation commode et valable seulement localement; c'est-à-dire, si les ordres de grandeurs des nombres mis en jeu sont comparables.

Le calcul des espérances offre, à mon avis, une bonne occasion de faire réfléchir à ce qu'est une approximation, et au domaine de validité d'un modèle.

c) ***Affaiblissement des hypothèses : Savage.***

Toute la construction, ici, a été, suivant la voie tracée par Pascal, fondée sur l'assimilation des paris à des vecteurs; cela a supposé au départ qu'un pari soit une liste de nombres: $x = (x_1, x_2, \ldots, x_n)$ et par conséquent, on y a insisté, que les conséquences de chaque éventualité i soient évaluées par un nombre x_i.

On peut refuser cette hypothèse dans certains problèmes de décisions face à l'incertitude; il y a bien des cas où les conséquences seront difficiles à évaluer numériquement dans une échelle d'utilité: par exemple, si des considérations de prestige, ou sentimentales, etc... entrent en jeu.

L'effort de L. J. Savage (voir (2) et (3)) a consisté à s'affranchir de l'hypothèse que les conséquences soient d'emblée évaluées numériquement; et à fonder son axiomatique sur la seule cohérence des décisions (de tenir ou non un pari) les unes par rapport aux autres. Le noeud de l'affaire, c'est bien entendu, chez lui aussi les décisions conditionnelles.

Et ce que montre Savage, c'est qu'un certain nombre d'exigences assez naturelles sur la cohérence (la "rationnalité") des décisions implique non seulement l'existence d'une distribution de probabilité (satisfaisant aux conditions habituelles) sur l'ensemble des éventualités, mais celle d'une utilité numérique sur les conséquences, et que l'utilité d'un pari est alors donnée par la règle de Pascal.

Bibliographie

1) Sur l'espérance comme utilité des paris, et le Calcul des décisions en cas d'incertitude.

(1) J. Von NEUMANN - Theory of games and Economic Behavior (P.U.P. - 1944) Chapitre 3.

(2) L.J. SAVAGE - The Foundations of Statistics (Wiley - 1954).

(3) G. MORLAT - Des poids et des choix. (Mathématiques et Sciences Humaines, n° 3 - 1963).

(4) G. MORLAT - Statistique et Théorie de la décision. (Mathématiques et Sciences Humaines, n° 8 - 1964).

(5) M. BARBUT - Calcul des décisions - Calcul des espérances - Calcul des Probabilités. (Mathématiques et Sciences Humaines, n° 2 - 1963).

2) Sur la règle des partis de Pascal.

(6) B. PASCAL - Lettres à Fermat du 29 Juillet et du 24 Août 1654. (Se trouvent dans les oeuvres complètes, Bibliothèque la Pléïade, N.R.F. et dans les "Oeuvres Complètes", L'intégrale, Edition du Seuil).

(7) G. Th. GUILBAUD - Eléments principaux de la théorie des jeux, in Stratégie et décisions économiques (C.N.R.S., 1954).

(8) G. Th. GUILBAUD - La règle des Partis et la Ruine des Joueurs. (Mathématiques et Sciences Humaines, n° 9 1964).

3) Un traité récent de Calcul des Probabilités (niveau 3ème Cycle) exploitent systématiquement la structure linéaire des espaces de variables aléatoires:

(9) J. NEVEU - Bases Mathématiques du Calcul des Probabilités - (Masson, 1964).