

La mode est aux mathématiques concrètes : bases de probabilités concrètes

par François PLUVINAGE, Strasbourg

1. Introduction

Cet article poursuit l'objectif ambitieux de présenter des évidences. L'objectif est ambitieux parce que le domaine concerné, les probabilités, passe pour exiger une certaine finesse, un certain "sens probabiliste". Même si la théorie apparaît comme simple à ses débuts (probabilités sur des ensembles finis), son application recèle de nombreux pièges. Et ensuite, c'est cette pilule des tribus.

Or, rien de tout cela n'est véritablement obligatoire. Pour ce qui est des probabilités sur des ensembles finis, la difficulté majeure me paraît être la notion d'événement, cette chose que l'on subit. Pourquoi ne pas la remplacer par l'expérience, cette chose que l'on provoque ? Au paragraphe 8.1., nous en verrons, sur un sujet de baccalauréat, un cas typique d'application. Les schémas d'expériences proposés sont des tirages dans des urnes, et une petite justification didactique s'impose, car tirer une boule dans une urne n'est ni palpitant, ni important. Mais tirer à la règle un trait sur une feuille de papier non plus. Ce qui compte dans un cas comme dans l'autre, c'est la neutralité du modèle, qui en autorise des applications extrêmement diverses, et sa proximité avec la théorie, qui permet dans bien des cas (et notamment dans tous les cas d'application courante) de penser de la même façon par référence au modèle ou à la théorie (au point de parfois les confondre). Ainsi dans l'enseignement, après une présentation de quelques minutes sur tous les domaines actuels où s'appliquent les probabilités, on pourra passer à une étude *contrôlable* avant d'aborder la spéculation pure : en effet, les machines programmables fournissent des possibilités de simulation de schémas d'urnes tout à fait convenables et d'un emploi très simple et très rapide (voir par exemple le paragraphe 8.2.). Pour les applications à des situations pratiques, on pourra se demander si un modèle d'urne donné est adéquat *avant* de lancer des calculs.

Pour ce qui est des tribus, il convient d'être plus nuancé. Ainsi, dans la brochure de l'A.P.M.E.P. intitulée *Hasardons-nous*, on peut trouver l'exposé de situations intéressantes dans un article de P.L. Hennequin : *Pourquoi des Tribus ?* Mais le même auteur indique dans un autre article (*Espérance mathématique et simulation*) que toute variable aléatoire peut être considérée comme variable aléatoire sur $]0,1[$ muni de la tribu des boréliens et de la mesure de Lebesgue. Ceci, compte tenu de l'article de Edward Nelson (*Internal Set Theory*, cité plus loin dans le texte), conduit bien à dire que toutes les probabilités se ramènent aux probabilités sur les ensembles finis. Pourvu, bien sûr, que l'on adopte le point de vue non-standard. Ce point de vue demande d'ajouter un prédicat (être standard), et trois axiomes concernant ce prédicat, aux axiomes ensemblistes usuels. Apparemment, ce n'est pas plus "méchant" que les tribus et la mesure, mais il s'agit en fait d'un changement dans nos habitudes de pensée, ce qui est toujours un peu délicat. Et puis nos Maîtres probabilistes n'ont pas fait comme ça pour établir leurs résultats. Mais justement, ils étaient des Maîtres... Digression : mon Maître G. Reeb m'a récemment signalé la tenue d'un congrès de probabilités non-standard au Japon, au mois de février 1980.

Sur le non-standard, je resterai donc réservé. Je serai moins réservé en revanche à propos de probabilités sur des ensembles finis. Il me paraît nécessaire que les premières constatations soient de véritables trivialisations, capables même d'engendrer le malaise diffus qui peut être ressenti par les débutants en géométrie (pourquoi nous dire ces évidences ?). Or seule l'équiprobabilité possède le caractère d'évidence souhaité. Je ne connais pas d'intuition probabiliste a priori qui n'ait son point de départ dans l'équiprobabilité. Ainsi, on sait ce qui se passe avec un dé, ou un autre polyèdre régulier homogène, mais on ne peut rien dire a priori sur la sortie de numéros qui auraient été placés sur un polyèdre biscornu (mais néanmoins assez "régulier" pour pouvoir rouler sur un plan). Le modèle de référence ne peut donc être qu'un modèle d'équiprobabilité. Le schéma d'urnes convient parfaitement bien et autorise tous les prolongements souhaitables. Notamment, l'urne de Bernoulli remplace très bien la notion d'événement : se poser une question sur la probabilité d'obtenir tel résultat, c'est lui substituer un tirage dans une unique urne contenant un certain nombre de boules blanches (correspondant à la réponse : Oui, on obtiendra ce résultat) et de boules noires (correspondant à la réponse : Non, on n'obtiendra pas ce résultat). Bien sûr, l'équiprobabilité ne débouche que sur des probabilités à valeurs rationnelles. Il y a donc lieu de prévoir des approximations pour les cas irrationnels. Ce n'est pas difficile d'imaginer cette extension, et en fait la pratique courante fait rencontrer peu de tels cas (d'ailleurs on calcule usuellement avec des décimaux n'ayant qu'un nombre limité de décimales). Alors, je ne vois pas d'inconvénient à "trivialiser" ainsi les probabilités élémentaires.

Remarquons que la situation est la même en géométrie élémentaire pour l'énoncé de Thalès : c'est un théorème pour les rapports rationnels,

un axiome pour les rapports réels quelconques. Nous reviendrons en conclusion sur le parallèle qu'il est possible d'établir entre géométrie et probabilités élémentaires.

2. Schéma d'urnes à boules individualisées

De même que l'on distingue axiomes et schémas d'axiomes, on peut parler d'expériences et de schémas d'expériences. Une expérience fixe les conditions, alors qu'un schéma laisse certains choix desquels les résultats peuvent dépendre.

A l'opposé des expériences déterministes ou des schémas d'expériences déterministes, un schéma d'expérience comme le suivant ne privilégie aucune issue parmi toutes celles qui sont possibles.

“Prendre une urne (ou un sac). Y placer des boules ayant toutes mêmes caractéristiques sensibles, mais munies toutes de numéros différents, de 1 à n s'il y a n boules. Agiter. Puis plonger la main dans l'urne (ou le sac) et, sans regarder, extraire une boule. Noter son numéro.”

A partir de ce schéma, on obtient une expérience en attribuant à n une valeur déterminée. Par exemple, on peut fixer $n=2$, et l'on peut dans ce cas substituer à notre expérience le jeu de pile ou face. Ou bien l'on peut fixer $n=6$ et jouer au dé. Ou encore $n=49$, comme au loto, où la procédure de tirage est très analogue à celle que nous avons décrite.

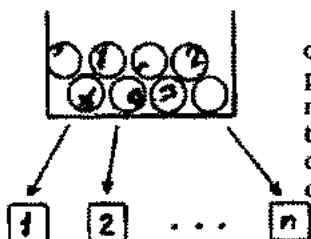
Remarquons que l'expérience peut aussi bien être reproduite par une même personne, qu'exécutée *simultanément* par plusieurs. Autrement dit, il n'y a pas d'ordre temporel, pas d'ordre algorithmique, entre plusieurs *réalisations* d'une expérience. Nous y reviendrons à propos d'indépendance.

Notre “expérience” (au sens courant de ce terme), traduite par ce que l'on nomme parfois le principe de raison insuffisante (aucun numéro n'a plus de raison qu'un autre de sortir), nous enseigne que chaque numéro, dans ces situations, sort autant que tout autre (statistique intuitive). Dans une situation de loterie équitable, chaque numéro donnera lieu à la même mise. Ainsi, s'il y a n numéros, chaque mise aura contribué dans la proportion $1/n$ à la somme totale misee. C'est pourquoi on a convenu que l'on attribuerait à chaque boule d'une expérience le “poids” $1/n$. Et, usuellement, ce poids prend le nom de probabilité.

Définition 1. Une structure finie d'équiprobabilité est la donnée d'un ensemble fini Ω et d'une même masse $1/n$ ($n = \text{card } \Omega$), encore appelée probabilité, attribuée à chaque élément de Ω .

Notation : $p(x)$ désignera la probabilité attribuée à $x \in \Omega$.

3. Partitions, plongements, espaces probabilisés finis



Comme le montre l'exemple de la loterie, on peut désigner *d'avance* les tirages éventuels, par exemple par des cartons. Mathématiquement, ceci revient à effectuer l'application identique sans changer les masses des éléments, ce qui conduit trivialement à la même structure d'équiprobabilité qu'au départ.

Le schéma d'expérience est toutefois modifié : le tirage effectué dans l'urne sert à désigner un tirage *noté d'avance*. Bien sûr, ceci suppose que tous les tirages possibles ont été représentés, chacun une fois.

Ainsi, dans une loterie, un même joueur peut miser sur plusieurs numéros. S'il mise sur p numéros parmi les n , sa mise représentera p/n de la mise totale. Ceci conduit à la règle d'additivité.

Définition 2. Soit une structure d'équiprobabilité finie, d'ensemble Ω . La masse, ou probabilité, attribuée à une partie A de Ω est par définition le produit de la probabilité de chaque élément par le nombre d'éléments de A .

Conséquences. On utilisera la notation $p(A)$ pour désigner la probabilité de $A \subset \Omega$.

3.1. Cas simples : $p(\emptyset) = 0$ et $p(\Omega) = 1$.

Si x est un élément quelconque de Ω , $p(\{x\}) = p(x)$; il n'y a donc pas de confusions possibles de notation dans l'emploi de p .

3.2. Soient A et B deux sous-ensembles quelconques de Ω . Il est immédiat que :

$$p(A \cup B) = p(A) + p(B) - p(A \cap B).$$

3.3. Si A_1, \dots, A_k est une partition de Ω en k sous-ensembles, alors :

$$\left\{ \begin{array}{l} \forall i, \forall j, (i \neq j) \Rightarrow p(A_i \cap A_j) = 0 \\ \sum_{i \in \{1, \dots, k\}} p(A_i) = 1 \end{array} \right.$$

Cas particulier : Partitions de la forme $\{A, \bar{A}\}$ où A est une partie propre de Ω .

3.4. Soit E un ensemble fini dont chaque élément x est muni d'une masse $P(x)$, avec :

$$\left\{ \begin{array}{l} \forall x \in E, P(x) \in \mathbb{Q} \quad \text{et} \quad 0 < P(x) < 1 \\ \sum_{x \in E} P(x) = 1 \end{array} \right.$$

La notation \mathbb{Q} désigne l'ensemble des nombres rationnels.

Alors il existe une structure d'équiprobabilité finie, d'ensemble Ω , une partition \mathcal{F} de Ω et une bijection $f: E \rightarrow \mathcal{F}$ vérifiant :

$$\forall x \in E \quad , \quad P(x) = p(f(x)) .$$

Démonstration : Représentons chaque $P(x)$ par une fraction, et soit a le dénominateur commun des fractions obtenues. Alors $a P(x)$ est un entier quel que soit x . Notons x_1, x_2, \dots, x_q les éléments de E et posons $k_1 = a P(x_1), \dots, k_q = a P(x_q)$. Les conditions voulues seront satisfaites par :

$$\Omega = \{1, 2, \dots, k_1 + k_2 + \dots + k_q\}$$

$$\mathcal{F} = \{A_1, A_2, \dots, A_q\}, \text{ avec}$$

$$A_1 = \{1, \dots, k_1\}, \dots, A_i = \{k_1 + \dots + k_{i-1} + 1, \dots, k_1 + \dots + k_{i-1} + k_i\}$$

$$f(x_i) = A_i \quad \text{pour tout } i = 1, \dots, q .$$

Définition 3. On appelle espace probabilisé fini la donnée d'un ensemble fini $E \neq \emptyset$ et d'une application $P: E \rightarrow]0;1[$ telle que

$$\sum_{x \in E} P(x) = 1 .$$

Terminologie : $P(x)$ est appelée la probabilité de l'élément x .

Remarque : Nous excluons ici l'éventualité d'éléments de probabilité nulle dans E .

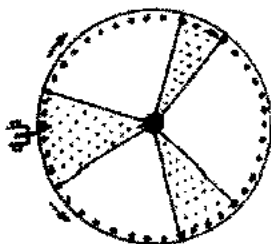
Proposition 1. Etant donné un nombre arbitraire ϵ , il existe une structure d'équiprobabilité finie, d'ensemble Ω , une partition $\mathcal{F} = \{A_1, \dots, A_q\}$ de Ω et une bijection $f: E \rightarrow \mathcal{F}$ telle que

$$\sup_{x \in E} |P(x) - p(f(x))| < \epsilon .$$

La démonstration est évidente : Il suffit de commencer à remplacer les $P(x)$ par des $P'(x)$ rationnels avec la précision voulue.

Terminologie : Nous dirons que tout espace probabilisé fini est, à une approximation arbitrairement fine près, *isomorphe* à une partition d'une structure finie d'équiprobabilité.

Réalisation expérimentale. Une roue de loterie découpée en secteurs fournit une bonne réalisation d'un espace probabilisé fini. Chaque secteur correspond à un élément x et $P(x) = \alpha/2\pi$, où α est l'angle en radians du secteur. Usuellement, on munit la roue, vers sa circonférence, de clous régulièrement espacés. A l'aiguille qui tourne (à moins que ce ne soit la roue elle-même qui tourne devant une marque fixe) est fixée une plaquette souple qui s'arrêtera dans un intervalle entre deux clous.



On peut espacer les clous avec une régularité parfaite si les $P(x)$ sont rationnels. Sinon, il sera possible de s'approcher d'une telle situation avec une précision arbitraire, limitée seulement par les éléments matériels (comme la grosseur des clous).

Proposition 2. Les propriétés 3.1., 3.2. et 3.3. sont valables pour tout espace probabilisé fini (E, P) .

La démonstration est évidente dans le cas rationnel. Sinon, il suffit de passer à la limite pour des approximations rationnelles de plus en plus fines.

4. Indépendance, structures d'équiprobabilité finies sur des ensembles produits

Considérons le même schéma d'expériences qu'au § 2, mais appliqué à deux urnes U_1 et U_2 , contenant l'une n_1 boules et l'autre n_2 boules. Comme les deux tirages, l'un dans U_1 et l'autre dans U_2 , peuvent être effectués dans un ordre arbitraire, la connaissance de l'un d'eux ne modifie en rien l'information sur le second — lequel reste soumis au principe de raison insuffisante. On peut aussi dire que tout couple de résultats peut sortir autant que tout autre.

Mathématiquement, nous déciderons qu'il y a équiprobabilité sur l'ensemble produit $\Omega_1 \times \Omega_2$, chaque fois que nous nous référons à un schéma d'urnes U_1 et U_2 avec ordre de tirages arbitraire (et éventuellement tirages simultanés dans chaque urne).

Il est facile de déterminer des algorithmes de tirages (chaînes de Markov par exemple) ne correspondant pas à cette situation. L'indépendance est donc essentiellement pour nous une condition de type algorithmique sur des expériences (plutôt que, comme on le dit usuellement, une contrainte sur des événements). Pratiquement, la question à se poser devant un cas d'application est de savoir si le modèle des deux urnes convient ou non.

Exemple : Tirages avec ou sans remise.

Si l'on considère deux tirages successifs dans une même urne, on peut ou non convenir de replacer dans l'urne la boule extraite au premier tirage. Si la boule est remise, le modèle d'indépendance est acceptable. Si la boule n'est pas remise, il n'est pas possible de substituer à l'urne pour le deuxième tirage une urne qui aurait été préparée à l'avance. L'équiprobabilité s'appliquera bien encore, mais à l'ensemble $\Omega \times \Omega - \{(x, x), x \in \Omega\}$, et non à $\Omega \times \Omega$ comme pour les tirages avec remise.

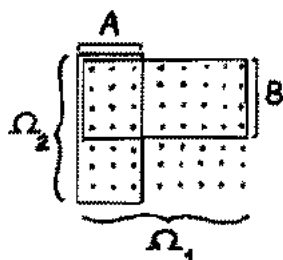
Proposition 3. Soit une structure d'équiprobabilité finie, d'ensemble $\Omega_1 \times \Omega_2$. Alors, quelle que soit $A_1 \subset \Omega_1$ et quelle que soit $A_2 \subset \Omega_2$

$$P_{\Omega_1 \times \Omega_2}(A_1 \times A_2) = P_{\Omega_1}(A_1) P_{\Omega_2}(A_2)$$

Note : Nous avons indicé les probabilités pour rappeler l'ensemble à considérer dans chaque cas. On peut souhaiter ne travailler que dans $\Omega_1 \times \Omega_2$, ce que permet le corollaire suivant (où ne figure pas l'indice qui serait $\Omega_1 \times \Omega_2$).

Corollaire. Sous les mêmes hypothèses que dans la proposition 3, la formule suivante régit les probabilités dans $\Omega_1 \times \Omega_2$.

$$p(A_1 \times A_2) = p((A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)) = p(A_1 \times \Omega_2) p(\Omega_1 \times A_2).$$



Si l'on pose $A_1 \times \Omega_2 = A$ et $\Omega_1 \times A_2 = B$, on se trouve ici devant ce qui est classiquement présenté comme des "événements indépendants" :

$$p(A \cap B) = p(A) p(B).$$

Nous dirons plutôt, lorsque cette relation est vérifiée dans un espace probabilisé fini (E, P) , qu'elle autorise à considérer pour les parties A et B que l'on se trouve dans la situation qui vient d'être décrite. C'est ce que dit la proposition suivante.

Proposition 4. Soit (E, P) un espace probabilisé fini, et soient A et B deux parties de E telles que $P(A) \in \mathbb{Q}$, $P(B) \in \mathbb{Q}$ et que :

$$P(A \cap B) = P(A) P(B).$$

Alors il existe une structure d'équiprobabilité finie sur un ensemble produit $\Omega = \Omega_1 \times \Omega_2$ et deux parties $A_1 \in \Omega_1$ et $A_2 \in \Omega_2$ telles que l'espace probabilisé fini $(\{A \cap B, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cap \bar{B}\}, p)$ soit isomorphe à $(\{A_1, \bar{A}_1\} \times \{A_2, \bar{A}_2\}, p)$.

Démonstration. Posons $P(A) = \alpha/a$ et $P(B) = \beta/b$, où α, a, β et b sont entiers. Nous prendrons alors pour Ω_1 un ensemble ayant a éléments et pour Ω_2 un ensemble ayant b éléments, puis pour $A_1 \subset \Omega_1$ un ensemble ayant α éléments et pour $A_2 \subset \Omega_2$ un ensemble ayant β éléments.

Nous avons bien $P(A \cap B) = P(A) P(B) = p(A_1 \times A_2)$.

$$\begin{aligned} \text{Et : } P(\bar{A} \cap B) &= P(B) - P(A \cap B) = P(B) (1 - P(A)) = p(A_1 \times \bar{A}_2) \\ P(A \cap \bar{B}) &= p(A_1 \times \bar{A}_2) \text{ de la même façon} \\ P(\bar{A} \cap \bar{B}) &= 1 - P(A \cap B) - P(A \cap \bar{B}) - P(\bar{A} \cap B) \\ &= p(\bar{A}_1 \times \bar{A}_2). \end{aligned}$$

On remarque que si $p(A)$ ou $p(B)$ n'avait pas été rationnel, on aurait pu s'approcher de cette situation aussi près qu'on l'aurait voulu.

5. Formule de Bayes

Partons du schéma d'expériences consistant en un tirage dans une urne contenant n boules, numérotées de 1 à n . Soit m un entier, $0 < m < n$. Décidons de ne noter, après tirage, un numéro que s'il est inférieur ou égal à m . Le principe de raison insuffisante qui gouvernait les tirages des boules de numéro dans $\{1, \dots, n\}$ s'applique aussi bien entre 1 et m . Autrement dit, ce nouveau schéma d'expériences correspond tout simplement à l'équiprobabilité sur $\{1, \dots, m\}$.

Proposition 5. Soit (E, p) un espace probabilisé fini, et soit F un sous-ensemble non vide de E . Alors $(F, \frac{1}{p(F)} p)$ est un espace probabilisé fini. En particulier, si (E, p) est une structure d'équiprobabilité, il en est de même de $(F, \frac{1}{p(F)} p)$.

La démonstration est évidente.

Terminologie usuelle : Pour $x \in F$, la probabilité $\frac{1}{p(F)} p(x)$ est nommée la probabilité de x conditionnellement à F , ou encore probabilité de x sachant F . On la note $p_F(x)$.

Extension : On convient généralement de définir p_F sur E tout entier. Il est alors naturel de poser $p_F(x) = 0$ si $x \notin F$. En effet, x est éliminé d'office si l'on rejette les résultats non dans F . Ainsi apparaît un cas où des éléments ont une probabilité nulle, ce que nous n'avions pas envisagé a priori (cf. § 3, déf. 3) ; c'est d'ailleurs pourquoi nous ne disons pas ici que (E, p_F) est un espace probabilisé fini. On peut en discuter ...

Propriétés des probabilités conditionnelles sur des espaces probabilisés finis.

5.1. Soit (E, p) un espace probabilisé fini, et soit F un sous-ensemble non vide de E . Alors :

$$\forall A \subseteq E, \quad p_F(A) = \frac{p(F \cap A)}{p(F)}.$$

C'est évident de par la définition de p_F .

5.2. Soient (E, p) un espace probabilisé fini, F et G deux sous-ensembles non vides de E . Alors :

$$p(G) = p_F(G) p(F) + p_{\bar{F}}(G) p(\bar{F}).$$

Démonstration. Cette égalité résulte immédiatement de 5.1. et de ce que

$$G = (G \cap F) \cup (G \cap \bar{F}).$$

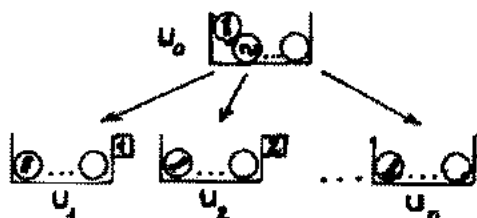
Proposition 6 (Formule de Bayes). Soient (E, p) un espace probabilisé, F et G deux sous-ensembles non vides de E . Alors :

$$p_G(F) = \frac{p_F(G) p(F)}{p_F(G) p(F) + p_{\bar{F}}(G) p(\bar{F})}$$

Démonstration. La proposition 6 est une conséquence immédiate de 5.1. et 5.2. .

6. Algorithmes de tirages

6.1. Tirage de Markov.



On considère le schéma d'expériences suivant. "Prendre une urne U_0 contenant n boules (comme au § 2), et n urnes numérotées de 1 à n contenant chacune des boules elles-mêmes numérotées. Effectuer un tirage dans U_0 puis un tirage dans l'urne U_k , où k désigne le numéro extrait de U_0 . Noter k et le numéro extrait de U_k ".

Ce que nous avons à notre disposition actuellement permet d'obtenir facilement la probabilité à attribuer à un élément quelconque (x, y) issu d'une telle expérience. En effet

$$p\{(x, y)\} = p\{(x, y) \cap U_x\} = p_{U_0}(x) p_{U_x}(y),$$

où p_{U_k} désigne l'équiprobabilité sur U_k , et où U_x dans $(x, y) \cap U_x$ désigne l'ensemble $\{(x, y), z \in U_x\}$.

Propriété. Si toutes les urnes U_1, \dots, U_n contiennent le même nombre m de boules, alors

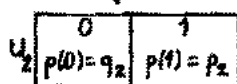
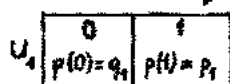
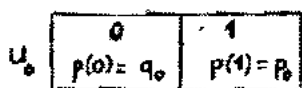
$$\forall y \in \{1, \dots, m\}, p\{1, \dots, n\} \times \{y\} = \frac{1}{m}.$$

Sous l'hypothèse de même composition des urnes U_1 à U_n , on est donc dans une situation d'indépendance : On pourrait substituer au schéma deux tirages indépendants dans les urnes U_0 et U_1 .

6.2. Application à des urnes de Bernoulli.

Une urne de Bernoulli correspond au cas particulier évoqué en 3.3. : Pour réaliser un espace probabilisé $\{0;1\}, p$ avec $p(1) = \hat{p} = \frac{a}{n}$ et

$p(0) = q = \frac{n-a}{n}$, on a recours à une urne contenant n boules dont a blanches (les "1") et $(n-a)$ noires (les "0").



On réalise le schéma d'expériences précédent avec trois urnes U_0, U_1 et U_2 de Bernoulli. Après tirage de 0 dans U_0 , on tirera dans U_1 , et après tirage de 1 dans U_0 , on tirera dans U_2 .

Note : Il s'agit bien du même schéma que précédemment, avec pour U_1, \dots, U_n des urnes de deux compositions, notées ici U_1 et U_2 .

Le tableau ci-contre indique les probabilités ainsi obtenues pour les couples

$$(x,y) \in \{0;1\} \times \{0;1\}.$$

Il est intéressant de former les *probabilités marginales* dans un tel cas. Evidemment, le premier tirage est effectué dans U_0 de composition donnée, et les probabilités p_0 et q_0 n'apprennent donc rien. Pour le second tirage, considéré seul (en oubliant le premier), nous

premier tirage deuxième tirage		0	1	
	0	$q_0 q_1$	$p_0 q_2$	$q_0 q_1 + p_0 q_2$
	1	$q_0 p_1$	$p_0 p_2$	$q_0 p_1 + p_0 p_2$
		q_0	p_0	

constatons que la situation est la même que celle d'une urne de Bernoulli telle que

$$\begin{cases} p(0) = q_0 q_1 + p_0 q_2 \\ p(1) = q_0 p_1 + p_0 p_2. \end{cases}$$

Il s'agit d'une situation barycentrique : Le second tirage considéré seul correspond à un tirage dans une urne de Bernoulli U' qui aurait une composition moyenne entre celle des urnes U_1 et U_2 . Exactement, cette urne pourrait être formée ainsi :

1° S'arranger pour que U_1 et U_2 contiennent autant de boules l'une que l'autre (réduction de p_1 et p_2 au même dénominateur).

2° Reproduire alors U_1 en autant d'exemplaires qu'il y a de boules 0 dans U_0 , et U_2 en autant d'exemplaires qu'il y a de boules 1 dans U_0 .

Former U' en mélangeant les contenus de toutes ces urnes.

Remarque : En mélangeant les contenus de deux urnes de Bernoulli quelconques, contenant l'une m boules dont a blanches et l'autre n boules

dont b blanches, on obtient la fameuse "règle d'addition" de fractions, trop familière à beaucoup de jeunes élèves

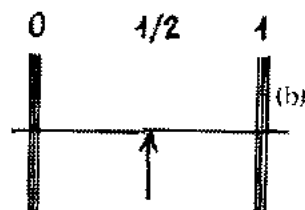
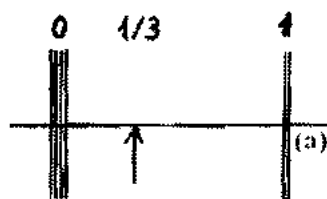
$$\frac{a}{m} \text{ " + " } \frac{b}{n} = \frac{a+b}{m+n}$$

Il est difficile, arrivé à ce point, de résister à l'envie de proposer d'autres exemples (urnes à plus de deux couleurs de boules, urnes de Polya, chaînes de Markov, ...). Mais cet article ne se veut qu'un exposé de principes, donc résistons-y ! Sauf pour ce qui est de la notion de variable aléatoire, dont l'importance justifie qu'elle ne soit pas passée sous silence ...

7. Variables aléatoires

De même que le schéma d'urnes nous a servi de support, de passerelle entre l'expérience concrète et la réflexion mathématique jusqu'à ce paragraphe, de même il est intéressant de disposer d'un nouveau support pour ce qui est de la notion de variable aléatoire. Celle-ci apparaît lorsque des résultats n'ont plus seulement comme jusqu'ici valeur qualitative, mais quantitative : jusqu'à présent, les numéros qui distinguaient deux boules dans une urne n'avaient qu'un rôle de désignation, donc la "différence" entre 2 et 3 par exemple était la même qu'entre 2 et 9, ou 1 et 100. Supposons à présent que les boules d'une urne désignent une quantité, une *grandeur*. Dans une situation de loterie, ceci est réalisé en convenant de gains différents selon les numéros ; par exemple, au jeu de dés nommé 421, les résultats correspondent à divers nombres de jetons.

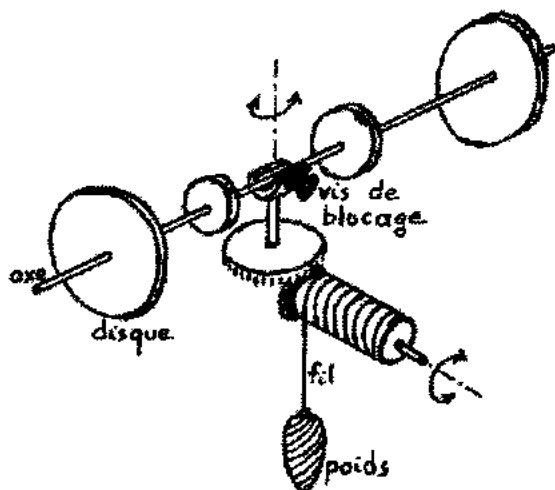
L'idée de représentation générale de cette situation d'application numérique définie sur Ω est à peu près évidente, puisqu'à chaque valeur de la variable est associée une probabilité. Il suffit d'une représentation plane, à deux axes dont l'un concerne les valeurs de la variable et le second leur probabilité. C'est le classique *diagramme en bâtons*, que l'on



peut aménager de bien des manières. Voici par exemple, sur les figures, un aménagement inspiré de barres d'haltères (à disques). La première figure (a) visualise $\{0;1\}$, $p(0) = 2/3$, $p(1) = 1/3$. La deuxième figure (b) visualise $\{0;1\}$, $p(0) = p(1) = 1/2$. Voir dans les axes de symétrie de ces figures les "barres d'haltères", et dans les segments verticaux les "disques d'haltères" vus par la tranche.

Pour permettre des comparaisons, nous avons chargé les deux barres de la même façon (6 disques au total). Au lieu de placer plusieurs disques identiques, on peut se donner des jeux de disques de masses 1, 2, 3, ... (pour les probabilités rationnelles — sinon il faudra se fabriquer des dis-

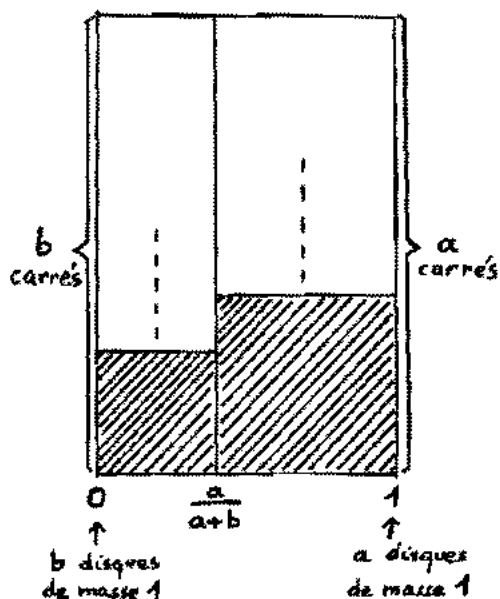
ques sur mesure), à placer sur un axe aux points d'abscisse convenable. Dans ce cas, il n'y aura qu'un disque en un point de l'axe. De toute façon, ici comme précédemment, on supposera la masse de la barre, constituant l'axe, négligeable par rapport à celle des disques.



Le dispositif ci-contre illustre les deux idées essentielles d'*espérance* et de *variance* d'une variable aléatoire sur un espace fini. L'*espérance* correspond à l'abscisse du centre de masse de l'"haltère" : C'est le point, matérialisé par la vis de blocage, où est réalisé l'équilibre. La *variance* correspond à l'inertie de l'"haltère". Cette inertie s'évalue grâce à l'accélération du poids

dans sa descente : l'observation de la durée de descente, à longueur de fil donnée et avec départ au repos, permet d'apprécier cette inertie. Montrer une fois un tel dispositif en fonctionnement ne me paraîtrait pas totalement inutile. Un peu de sens mécanique par exemple fait deviner que, parmi les deux figures qui précèdent, la première correspond à un système un peu moins inerte que le second, les centres de masse étant désignés par les flèches.

Si l'on veut, il est possible de proposer pour l'inertie, autrement dit la variance, une construction purement géométrique. Ainsi une manipulation est possible uniquement avec papier-crayon pour matériel. L'idée d'une construction géométrique est naturelle, puisqu'un moment d'inertie est le produit d'une masse par le *carré* d'une distance : il suffit de dessiner des carrés. A titre d'exemple, traitons le cas binomial, avec $p(1) = p$ et $p(0) = q = 1 - p$, en supposant p rationnel (comme déjà dit, les cas irrationnels font l'objet d'approximations). Donc, on peut poser :



$$p = \frac{a}{a+b} \text{ et } q = \frac{b}{a+b},$$

avec a et b entiers.

La "barre d'haltères" correspondante possède alors une inertie centrale qu'illustre la figure ci-contre, par l'aire du rectangle obtenu en accolant des carrés construits sur les extrémités 0 ou 1 et le centre de masse $\frac{a}{a+b}$. La

variance ne sera autre que l'aire moyenne des carrés, c'est-à-dire, puisqu'il y a $(a+b)$ carrés :

$$\frac{1}{a+b} \left(b \times \left(\frac{a}{a+b} \right)^2 + a \times \left(\frac{b}{a+b} \right)^2 \right) = \frac{ab}{(a+b)^2} = pq.$$

8. Applications

8.1. Un sujet de baccalauréat.

Le sujet donné au baccalauréat à Paris en 1975 me paraît intéressant pour illustrer la simplicité du point de vue expérimental, ou algorithmique, par rapport au concept d'événements dans un espace probabilisé. Ce problème a fait l'objet de commentaires dans le Bulletin de l'A.P.M.E.P. n° 302 (février 1976). La Régionale de Paris qualifie le problème III d'astucieux et A. Torrat dit que sa motivation probabiliste déborde le cadre précis du programme. C'est à ce propos que je voudrais intervenir, car par ailleurs je n'ai évidemment rien à redire sur les critiques concernant la longueur du sujet complet. Mon propos est de montrer que le problème III est enfantin traité en schéma d'urnes, alors qu'il demande de la finesse probabiliste si l'on veut suivre une démarche reposant sur les événements. Voici l'énoncé de ce problème, qu'il convient d'accompagner de l'indication : Dans la partie II de ce sujet, non reproduite ici, on étudiait l'espace des suites u_n satisfaisant à la relation de récurrence

$$\forall n \in \mathbb{N} \quad , \quad p u_{n+2} - u_{n+1} + (1-p) u_n = 0.$$

Notamment, il s'agissait d'obtenir dans cette partie II l'expression de u_n sous l'hypothèse que $u_0 = 1$ et $u_\alpha = 0$ (α entier positif donné).

III. Un jeu oppose deux joueurs A et A', auxquels on attribue respectivement, au début du jeu, un "avoir" de a jetons et un "avoir" de 2a jetons (a entier donné, supérieur ou égal à 1). La rencontre comporte des parties successives et indépendantes, numérotées 1, 2, 3 ...

La probabilité pour que le joueur A gagne une partie est supposée indépendante du rang de cette partie, et égale à p ($0 < p < 1$). Après chaque partie le joueur perdant donne un jeton au gagnant. Le jeu s'arrête lorsqu'un joueur est "ruiné", c'est-à-dire ne dispose plus de jetons, et le joueur "ruiné" perd le match.

1. a) k désignant un entier naturel, on considère la variable aléatoire X_k égale à l'avoir du joueur A après la partie de rang k si $k \neq 0$ et avant la partie de rang $k + 1$ (si celle-ci a lieu). On a ainsi $X_0 = a$ et $0 \leq X_k \leq 3a$.

Quelles sont les valeurs "possibles" de X_1 ? de X_2 ? de X_{2k} ? de X_{2k+1} ?

b) Si $X_k = 0$ le joueur A est ruiné ; si $X_k = 3a$ le joueur A' est ruiné ; dans chacun de ces cas le match ne se poursuit pas au-delà de la k^{ème} partie.

Si X_k est différent de 0 et de 3a, l'on admet (*) que la probabilité de ruine ultérieure du joueur A ne dépend pas de k mais seulement de la valeur n de X_k .

On désigne par r_n la probabilité de ruine de A, connaissant n. On a ainsi $r_0 = 1$ et $r_{3a} = 0$.

En considérant les deux valeurs que peut prendre X_{k+1} sachant que $X_k = n$, montrer (**) que $r_n = (1-p)r_{n-1} + pr_{n+1}$, et constater que la suite $n \rightarrow r_n$ vérifie la relation de récurrence (1) du II.

Exprimer alors, à l'aide de II. 2. b), le terme r_n en fonction de n et de a (lorsque $p = \frac{1}{2}$) ou en fonction de n, a et de $x = \frac{1-p}{p}$

(lorsque $p \neq \frac{1}{2}$).

c) On désigne par r'_m la probabilité de ruine du joueur A', connaissance son avoir m.

Montrer qu'on obtient r'_m en remplaçant, dans l'expression de r_n , n par m et p par $1-p$ (c'est-à-dire x par $\frac{1}{x}$).

(*) Le candidat ne cherchera pas à définir l'espace de probabilité relatif à ce jeu, et se bornera à faire le raisonnement qui lui est suggéré.

(**) Le candidat pourra admettre ce résultat.

Ecrire cette expression de r'_m (pour $p = \frac{1}{2}$ et pour $p \neq \frac{1}{2}$).

Vérifier la relation $r_a + r'_{2a} = 1$ (2).

2. En notant que r_a et r'_{2a} sont les probabilités de ruine de A et de A' au début du match, on voit que le jeu est favorable au joueur A si $r_a < r'_{2a}$, c'est-à-dire, d'après la relation (2) précédente, si $2r_a < 1$.

Que vaut r_a lorsque $p = \frac{1}{2}$?

On prend $p \neq \frac{1}{2}$. Exprimer la différence $D_a = 2r_a - 1$ en fonction de x et de a .

Pour quelles valeurs de x a-t-on $D_a < 0$? (cf. le 1. 1.).

p étant fixé, supérieur à $\frac{1}{2}$, comment choisir a pour que le jeu soit favorable au joueur A ?

Application numérique : $p = 0,51$; utiliser le 1. 2. pour donner la plus petite valeur convenable de l'entier a .

Remarquons les difficultés indiquées par les auteurs du sujet eux-mêmes, dans les deux petites notes de bas de page. Et quel peut bien être le raisonnement invoqué dans la première de ces deux notes ? A mon avis, ce n'est rien d'autre qu'une référence implicite aux schémas d'urnes, référence que je vais maintenant expliciter (ou du moins essayer, si le lecteur veut bien me suivre).

La question de la ruine du joueur A est la suivante, sous forme de schéma d'urne :

"Peut-on remplacer le jeu tel que l'énoncé le définit par un tirage dans une urne contenant deux sortes de boules : les boules 0 indiquant que A est ruiné et les boules 1 indiquant que A n'est pas ruiné ?"

Evidemment, la réponse est non. Ce schéma n'est pas satisfaisant parce que l'avoir n ($0 < n < 3a$) de A n'est pas pris en compte.

Améliorons la question.

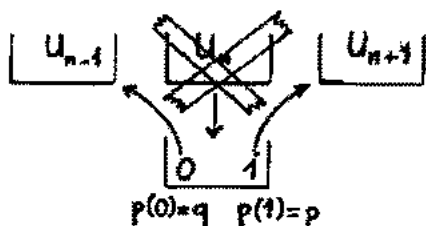
"Peut-on remplacer le jeu tel que l'énoncé le définit par un tirage dans l'une de $(3a + 1)$ urnes (U_0, U_1, \dots, U_{3a}) contenant chacune deux sortes de boules : les boules 0 et 1 (comme précédemment 0 signifie que A est ruiné) ?"

Ainsi posée, la question est raisonnable, car l'indépendance invoquée dans l'énoncé permet d'accepter que seul l'avoir de A, et non le numéro du coup du jeu, entre en ligne de compte. Mais des conditions apparaissent, qui sont les suivantes :

1° U_0 ne contient que des boules 0.

2° U_{3a} ne contient que des boules 1.

3° Pour tout n , $0 < n < 3a$, un tirage dans U_n peut être remplacé par un coup de jeu suivi d'un tirage dans U_{n-1} (si A perd) ou dans U_{n+1} (si A gagne).



Il s'agit donc très exactement du problème de double tirage dont nous avons parlé en 6.2. L'urne U_n n'est autre que l'urne "marginale" U' qui peut être substituée à U_{n-1} et U_{n+1} . La relation

$$r_n = qr_{n-1} + pr_{n+1},$$

où r_k désigne la proportion de boules 0 dans U_k , est donc immédiate après l'étude générale, proposée au § 6. N'est-ce pas un tout petit peu plus simple que d'envisager $\cup(\{0;1\}^K, K \in \mathbb{N})$?

Le reste du problème n'offre pas de difficulté de résolution. Il y a simplement la surprise de trouver effectivement la relation

$$r_a + r_{2a} = 1.$$

Ainsi notre jeu aboutira "toujours" à la ruine d'un des deux joueurs. Les "événements de probabilité nulle" pourront donc surgir ici à propos des parties possibles à ce jeu, mais après coup et non a priori.

8.2. Simulation sur machines

En soi, la notion de "suite aléatoire" est contradictoire avec le schéma de tirage dans une urne. Un excellent article de C. Dellacherie montre qu'en définitive on ne peut parler que de suites illimitées *plus ou moins* aléatoires. En effet, la possibilité de tirages simultanés s'oppose à toute idée d'algorithme de la forme $u_n = f(u_{n-1})$. Cependant de tels algorithmes peuvent être susceptibles de résister à la découverte précisément de leur caractère d'algorithmes ; il peut ainsi arriver que, du point de vue de la *prédiction de résultats* par un observateur humain, ils puissent être aussi "aléatoires" (ou presque) qu'un tirage dans une urne. Le "bon" critère à envisager est d'imaginer une situation de test, du type séries à compléter ; l'observateur humain est censé ignorer f . L'algorithme lui fournit des résultats. On dira que l'algorithme *résiste* tant que l'observateur ne réussit pas à prévoir le résultat à venir, mieux qu'un résultat de tirage dans une urne. Pour atténuer le caractère en apparence subjectif de ceci, on peut supposer que l'observateur dispose de l'arsenal statistique et probabiliste. Notre observateur humain n'est donc pas vous ou moi, mais "on". Ainsi, "on" pourra commencer à avoir une petite idée en voyant :

1 ...

ou

1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 ...

ou

0 1 0 0 1 1 0 0 0 1 1 1 0 0 0 0 1 1 1 1 0 0 0 0 0 ...

En revanche, ce ne sera pas le cas devant :

0 0 1 1 1 0 0 0 1 1 0 1 1 1 1 1 0 0 0 0 0 0 1 1 0 ...

Même en continuant longtemps à "tirer" des numéros 0 ou 1, "on" aura du mal à mettre en avant un algorithme équivalent à celui qui est utilisé ici. Cet algorithme est le suivant, en désignant par $\{x\}$ la partie entière de x

$$v_n = [2u_n], \text{ avec } \begin{cases} u_0 = 0,13579 \\ u_{n+1} = 997 \times u_n - [997 \times u_n] \end{cases}$$

Une suite v_n ainsi obtenue est nommée *pseudo-aléatoire*. Ici, notre suite sera forcément périodique, sa période ne pouvant pas avoir une longueur de plus de 10000 (puisque u_n a exactement 5 chiffres significatifs). Donc "on" ne pourra manquer de la "démasquer" après 10000 tirages au plus.

En partant de $u_0 = 0,5284163$ on obtiendrait une suite dont la période a une longueur voisine de 500 000, et "on", disposant des tests connus, ne parviendrait pas à la "démasquer" avant un nombre considérable de tirages. En effet, en extrayant de u_n ses cinq premiers chiffres significatifs, on obtient des résultats qui simulent très bien un tirage aléatoire avec équiprobabilité dans $D_5 \cap]0,1[$ (décimaux à 5 chiffres de]0,1[). De ce fait, on peut en déduire de multiples simulations de tirages aléatoires variés. Par exemple :

1° $v_n = [au_n]$, avec $1 < a < 2$, simule une urne de Bernoulli avec $p(0) = 1/a$ et $p(1) = (a-1)/a$.

2° $v_n = 1 - [au_n]$, avec $1 < a < 2$, simule une urne de Bernoulli avec $p(0) = (a-1)/a$ et $p(1) = 1/a$.

Ainsi, on peut simuler une urne de Bernoulli arbitraire et, avec une machine programmable, il est facile également d'organiser des simulations de tirages de Markov. En prenant $a=6$ au lieu de $1 < a < 2$ dans le premier cas ci-dessus, on "jouera au dé". Bref, les possibilités d'organiser des expériences multiples avec facilité et rapidité sont très abondantes. Il n'est pas exagéré de dire que nous disposons dans ce domaine de possibilités expérimentales comparables au dessin de figures en géométrie.

9. Conclusion : Un parallèle avec l'enseignement des concepts numériques et celui de la géométrie élémentaire

Dans les deux domaines déterministes de base, que sont le calcul arithmétique et la géométrie élémentaire, on peut séparer deux temps d'appropriation par un individu. Le *premier temps* pourrait être nommé

celui de la *mise à l'épreuve* : une notion (par exemple la somme, le produit, le parallélisme, ...) est utilisée dans des situations où une procédure ne la mettant pas en œuvre est également possible. Ainsi, on peut substituer à une addition simple un comptage, sur les doigts par exemple. Ou bien l'on peut substituer une mesure de longueur à un report par parallélisme. Le *deuxième temps* est celui de *l'utilisation en confiance*. Par exemple, effectuation de $1347 + 694$ ou mesure de la distance Terre-Lune grâce à des repérages angulaires.

Certainement, des décalages existent entre diverses notions de ces deux domaines. Ainsi la multiplication peut-elle être mise à l'épreuve par l'addition. Mais il faut avoir initialisé le processus, amorcé la possibilité d'appropriation.

Dans un domaine très neuf pour les élèves, prévoir de façon substantielle cette initialisation n'est sûrement pas superflu. C'est un peu ce qu'il faut voir dans cet article sur les débuts d'un apprentissage probabiliste.

Bibliographie

- C. Dellacherie, *Nombres au hasard*, Gazette des mathématiciens n° 11, octobre 1978.
- A. Engel, *Mathématiques élémentaires d'un point de vue algorithmique*, CEDIC, Paris, 1978.
- P.L. Hennequin, in *Hasardons-nous*, Brochure A.P.M.E.P. n° 17, 1976.
- E. Nelson, *Internal Set Theory*, Bulletin of the American Mathematical Society, Vol. 83, number 6, novembre 1977.
- A. Tortrat, *Sur le problème de TC*, Paris, juin 1975, Bulletin de l'A.P.M.E.P. n° 302, février 1976.