Daniel Saada Conférence APMEP du 23/10/05 dsaada@yahoo.fr Lycée Hoche, Versailles

Les Fondements du Calcul des Probabilités

0) Une « expérience aléatoire » produit des résultats dont l'ensemble, appelé univers des possibles, est noté traditionnellement Ω. Nous ne parlerons ici que d'univers infinis.

Exemple **:** *le jeu infini de Pile ou Face.*

On lance indéfiniment une pièce de monnaie : si on note 1 l'apparition de pile et 0 l'apparition de face, l'univers Ω de tous les résultats est $\{0,1\}^{\mathbb{N}^*}$, Ω ayant la puissance du continu.

En général, les probabilités des événements élémentaires sont soit dictées par l'expérience soit imposées par l'utilisateur, le Calcul des Probabilités se chargeant d'évaluer les autres probabilités ; dans le jeu *infini de Pile ou Face*, on fixe évidemment à 0 les probabilités élémentaires ($1/2^n$ si on lance la pièce n fois).

1) Le Théorème d'Ulam (1930) et la fin de l'Eden probabiliste

Si Ω est un ensemble infini équipotent à R , il n'existe pas de probabilité définie sur P(Ω) tout entier qui ne charge aucun point de Ω (on dit que p est diffuse).

On prouve en effet que si p(ω) est nul pour tout ω de Ω, alors p(Ω) est nul ; en d'autres termes, la fonction d'ensemble p, qui s'annule par hypothèse sur les parties dénombrables de Ω, vaut aussi 0 sur toutes les parties de cardinal $\aleph_1$: un grand mystère pour une fonction σ-additive ! On pourrait dire aussi que p est $\aleph_1$-additive.

La clé est la suivante : Si un ensemble Ω a la puissance du continu, on peut le munir d'une relation d'ordre total pour laquelle le sous-ensemble des x < ω est dénombrable, pour tout ω de Ω.

Cette clé est une conséquence du **Théorème de Zermelo (1904) : tout ensemble peut être muni d'un bon ordre,** un ordre pour lequel toute partie non vide de E contient un plus petit élément.

Pour une démonstration, voir **[5]** chapitre **II**, le point essentiel étant que ce théorème **est équivalent aussi bien à l'axiome du choix qu'au théorème de Zorn (1935, mais Kuratowski en 1922).**

Th. de Zorn : *Si E est un ensemble ordonné dont toutes les parties totalement ordonnées sont majorées, alors E a un élément maximal.*

Une application : tout espace vectoriel a une base, car l'ensemble des parties libres est *inductif* .

Conclusion : si p ne charge aucun point de Ω, p ne pourra être définie que sur « une sous-tribu » stricte de Ω, **et donc se résigner au fait troublant qu'il y aura nécessairement des parties sans probabilités, des événements non probabilisés.** C'est le cas du *jeu infini de Pile ou Face*.

Consolation ? Il existe sur P(Ω) des probabilités si on abandonne la condition p(ω) nulle pour tout ω de Ω, mais elles sont très pauvres : ce sont les probabilités dites *discrètes*.

démonstration : **a)** la partie Ω' formée des ω de probabilité non nulle est au plus dénombrable

car si $\Omega'_n = \{\omega : p(\omega) \geq 1/n\}$, alors Ω' = $\bigcup_{n\geq 1}\Omega'_n$ et card $\Omega'_n \leq$ n.

b) p(Ω') = 1 en vertu du th. d'Ulam

Ω - Ω' a encore la puissance du continu ; comme p est définie aussi sur $\mathcal{P}$ (Ω - Ω') et ne charge aucun point de Ω - Ω', p(Ω - Ω') = 0, et donc p(Ω') = 1.

c) p est discrète : pour toute partie A de Ω, p(A) = p(A ∩ Ω') + p(A ∩ (Ω - Ω')) = p(A ∩ Ω') ; p est donc discrète, l'univers « réel » étant la partie dénombrable Ω'.

2) Probabilités sur un univers dénombrable Ω

Les probabilités p sur P(Ω) sont parfaitement identifiées : on se donne une série à terme général $p_n \geq 0$, de somme 1, on décrète que $p(\omega_n) = p_n$, et on pose, pour tout A dans Ω, p(A) = $\sum_{\omega_n \in A} p(\omega_n)$,

On démontre que toute probabilité sur un univers fini ou dénombrable **se prolonge** à $\mathcal{P}$ (Ω) tout entier.

Donc, <u>la notion de tribu est superfétatoire pour de tels univers.</u>

Deux exercices difficiles sur les probabilités dénombrables (on choisit Ω = N par commodité) :

i) <u>Limite d'une suite de probabilités</u> : on suppose que pour tout k, $p_n(k)$ a une limite, notée p(k).

a) Il se peut que $\sum_N p(k) < 1$ et même que $\sum_N p(k) = 0$.

exemple :on tire n fois sans remise dans {1, 2,..., 2n}, la probabilité d'obtenir k entiers pairs est $p_n(k) = (C_n^k)^2 / C_{2n}^n$.
Or, pour k fixé, la limite de $p_n(k)$ pour n infini est 0 car $p_{n+1}(k)/p_n(k) \to 1/4$; p est donc la fonction nulle !

b) On suppose désormais $\sum_N p(k) = 1$: p est alors une probabilité sur N, définie par p(A) = $\sum_{k \in A} p(k)$.

Peut-on affirmer que p(A) est la limite des $p_n(A)$? La réponse est oui.

On regarde chaque p_n <u>comme une application sur R^+</u> en posant $p_n(x) = p_n(k)$ si x est dans [k,k+1[.

On a alors : **a)** $p_n \to p$ **b)** p_n et p ≥ 0 **c)** les intégrales de 0 à +∞ de p_n et p valent 1.

En appliquant le Théorème de convergence dominée de Lebesgue à la suite $f_n = \min(p,p_n)$, on obtient en effet que $p_n(A) \to p(A)$, et ceci uniformément en A.

ii) <u>L'ensemble des valeurs prises par toute probabilité p sur N est un fermé de</u> [0,1] .

L'adjectif fermé signifie que si $p(A_n)$ a une limite ℓ, alors il existe A dans Ω tel que p(A) = ℓ.

L'idée cette fois est d'écrire p(A) = $\sum_{n \in A} p_n$ sous la forme $\sum_n 1_A(n) p_n$ et de considérer p comme une application définie sur E = $\{0,1\}^N$ par : si s est une suite de E, p(s) = $\sum_0^\infty s_n . p_n$

On munit E de la distance d(s,t) = $\sum_0^\infty \frac{|s_n - t_n|}{2^{n+1}}$: E est un espace <u>compact</u> pour cette distance.

On prouve enfin que p est *continu* sur E : p(E) est donc un compact de [0,1], donc est un fermé.

3) Construction de probabilités

Reprenons le *jeu infini de Pile ou Face* : parmi les événements de $\{0,1\}^{N^*}$, quels sont ceux que l'on peut probabiliser ? Et comment ?

Autre exemple fondamental : longueur d'une partie de [0,1].

Soit J un intervalle (a,b) de l'univers Ω = [0,1] : on pose p(J) = b – a. En dehors des intervalles, quelles sont les parties de [0,1] que l'on va pouvoir mesurer ? Pas toutes d'après le Théorème d'Ulam puisque, par construction, la longueur de tout point est 0.

Voici la démarche commune pour solutionner ces deux problèmes.

Étape 1 : on étend p à une algèbre d'événements

pour les longueurs : l'ensemble $\mathcal{A}$ des réunions finies d'intervalles disjoints est une algèbre car

- $\mathcal{A}$ est stable par complémentaire (un dessin aide à formaliser),
- $\mathcal{A}$ est stable par intersection en vertu de la formule ensembliste $\sum_1^n I_p \cap \sum_1^m J_q = \sum_{p,q} I_p \cap J_q$, le symbole $\sum$ désignant les réunions disjointes.

Pour A = $\sum_1^n (a_i, b_i)$, on pose p(A) = $\sum_1^n (b_i - a_i)$, mais il faut s'assurer au préalable que p(A) ne dépend pas de la réunion choisie.

pour Pile ou Face : soit P_n l'ensemble des suites s de $\{0,1\}^{N^*}$ telles que $s_n = 1$, et F_n le complémentaire de P_n.
Evidemment encore, $p(P_n) = p(F_n) = ½$ pour tout n ; les intersections finies de P_i et/ou F_j ont pour probabilité $(1/2)^k$, k étant le nombre de facteurs de l'intersection, pour respecter l'indépendance des lancers.
Les réunions finies disjointes de ces intersections forment une algèbre : on les probabilise en additionnant les probabilités des intersections.

Étape 2 : on vérifie que p est σ-additive sur l'algèbre construite à l'étape 1

pour les longueurs, la difficulté est en amont, et réside en ceci : si (a,b) = $\sum_1^\infty (a_n, b_n)$, alors b – a = $\sum_n (b_n - a_n)$.

La démonstration n'est en rien anodine puisqu'elle utilise la **compacité** des segments de R : on « minore » (a,b) par le compact [a +ε, b -ε] et on « majore » (a_n, b_n) par l'ouvert $]a_n - \varepsilon/2^n, b_n + \varepsilon/2^n[$.

Étape 3 : on finit avec Le Théorème de **Carathéodory** (1927) qui affirme

Toute probabilité σ-additive définie sur une algèbre se prolonge
à la tribu engendrée par l'algèbre et ce prolongement est unique.

C'est l'existence et l'unicité qui permettent d'attribuer une probabilité
à des événements qui n'étaient pas initialement dans l'algèbre.

Par exemple, on démontre que l'ensemble C des suites s telles que $\frac{1}{n}\sum_1^n s_i$ tend vers ½ est dans la tribu engendrée par les P_n et les F_n, puis que p(C) = 1 : **c'est l'énoncé de la loi forte des grands nombres**.
La tribu engendrée par l'algèbre $\mathcal{A}$ s'appelle la tribu de **Borel.** En vertu du Th. d'Ulam, il existe des non boréliens ; il en existe même infiniment, car la tribu de Borel n'a que la puissance du continu.

4) De Borel à Lebesgue

Soit p une probabilité sur une tribu T d'un univers Ω : un sous-ensemble N de Ω est dit négligeable pour p si N est contenu dans une partie B de la tribu de probabilité nulle : $N \subset B$, $B \in \mathcal{T}$ et p(B) = 0.
On démontre sans difficulté que :

1) L'ensemble des $A \cup N$, avec A dans $\mathcal{T}$ et N négligeable, est une tribu, notée $\widehat{\mathcal{T}}$, contenant $\mathcal{T}$.

2) La probabilité p se prolonge en une probabilité sur $\widehat{\mathcal{T}}$: c'est l'extension ultime de p.

Quand $\mathcal{T}$ est la tribu de Borel, $\widehat{\mathcal{T}}$ est la tribu des mesurables de Lebesgue :
ainsi passe-t-on des boréliens aux mesurables, et des Probabilités à l'Intégration.

Bibliographie

A) Les livres qui m'ont aidé à bâtir cet exposé (extrait d'un livre à venir)

[1] Paul Louis Hennequin, *Pourquoi des tribus ?*, brochure n° 17 de l'APMEP, 1976.

[2] Jacques Neveu, *Bases mathématiques du calcul des probabilités*, Masson, 1964.

[3] O. Arino, C. Delode, J. Genet, *Mesure et Intégration, exercices et problèmes*, Vuibert, 1976.

[4] Dominique Foata et Aimé Fuchs, *Calcul des Probabilités*, Masson, 1996.

[5] Gustave Choquet, *Outils topologiques et métriques de l'Analyse mathématique*,

cours rédigé par Claude Mayer, Centre de Documentation Universitaire, 1969.

[6] Paul Deheuvels, *La probabilité, le hasard et la certitude*, Que sais-je ?, PUF, 1990.

[7] Jean-Marie Arnaudies, *L'intégrale de Lebesgue sur la droite,* Vuibert, 1997.

[8] Jean Genet, *Mesure et Intégration,* Vuibert, 1976.

B) Les articles que j'ai écrit pour le Bulletin de l'APMEP :

Une liste figure à l'adresse :
http://publimath.irem.univ-mrs.fr/cgi-bin/publimath.pl?r=auteur%3D%22Saada+Daniel%22

le dernier article, portant sur la loi binomiale, est en ligne à :
http://lerouge.univ-lyon1.fr/apmep/spip/IMG/pdf/BV436LMB.pdf

ou, plus simplement, taper « Daniel Saada » dans un moteur de recherche.