Espoir

Karim Zayana[1,2] et Victor Rabiet[1,3]

[1] Ministère de l'Éducation nationale, Paris
[2] LTCI, Télécom Paris, Institut Polytechnique de Paris
[3] DMA, École normale supérieure, Paris

En classes de Première et Terminale [6, 7], le lycéen croisera parfois deux formules de l'espérance d'une variable aléatoire (discrète finie) X. Voyons cela quand

$$X : \Omega = \{\omega_1, \omega_2, \ldots, \omega_n\} \to X(\Omega) = \{x^1, x^2, \ldots, x^p\} \subset \mathbb{R}$$

où, pour différencier les éléments de l'univers Ω de ceux du support $X(\Omega)$, nous avons énuméré les uns par des indices i, les autres des exposants j.

La première formule moyenne *sans répétition* les *images* x^1, x^2, …, x^p de X en les pondérant de leurs probabilités d'être atteintes. Soit

$$\mathrm{E}[X] = \sum_{j=1}^{p} x^j \mathbb{P}_X([X = x^j]) \tag{1}$$

où $\mathbb{P}_X([X = x^j])$, défini par $\mathbb{P}(\omega \in \Omega \,/\, X(\omega) = x^j)$, correspond à la *probabilité image* de $\mathbb{P}$ par X. L'expression (1), « pilotée par l'*output* », adopte une logique de sommation *en tranches* dans l'esprit de l'intégrale de Lebesgue [1]. C'est la définition de l'espérance dont partent en général les statisticiens [1, 2]. Pour intuitive qu'elle soit, elle en occulte une propriété considérable, la linéarité, illisible sinon au prix de vilaines contorsions.

La deuxième formule « rétro transfère » le calcul de $\mathrm{E}[X]$ à l'univers Ω des issues. La somme obtenue procède alors d'une réorganisation de (1) que permettent la commutativité et l'associativité de l'addition. Elle moyenne de fait *avec répétition* les images de X par les fréquences des événements élémentaires, adoptant une logique de sommation en piles (un peu « à la Riemann [2] », pour filer l'analogie [3] de l'intégration), pilotée par l'*input* :

$$\mathrm{E}[X] = \sum_{i=1}^{n} X(\omega_i)\mathbb{P}(\omega_i) \tag{2}$$

Cette deuxième écriture, plus formelle, est celle des probabilistes [3, 4, 5, 8]. Pour être tout à fait honnête, il s'agit de la « vraie » définition de l'espérance, dont tout découle ensuite. En particulier que pour deux variables aléatoires X et Y, $\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$, et ce sans avoir eu à déterminer la loi de $X+Y$. Le mécanisme du transfert de (2) vers (1) peut être appréhendé sur des exemples génériques, en les accompagnant d'une illustration. Proposons celle-ci, où les $n = 6$ germes de probabilité sont figurées par de petits poids de valeur totale l'unité de masse et les $p = 4$ images par des petites pastilles.

1. Henri LEBESGUE ; mathématicien français, 1875 – 1941.
2. Bernhard RIEMANN ; mathématicien allemand, 1826 – 1866.
3. Il ne faut pas prendre au pied de la lettre ces deux analogies : si l'intégrale de Lebesgue peut, elle, admettre une version discrète, celle de Riemann n'en a pas.

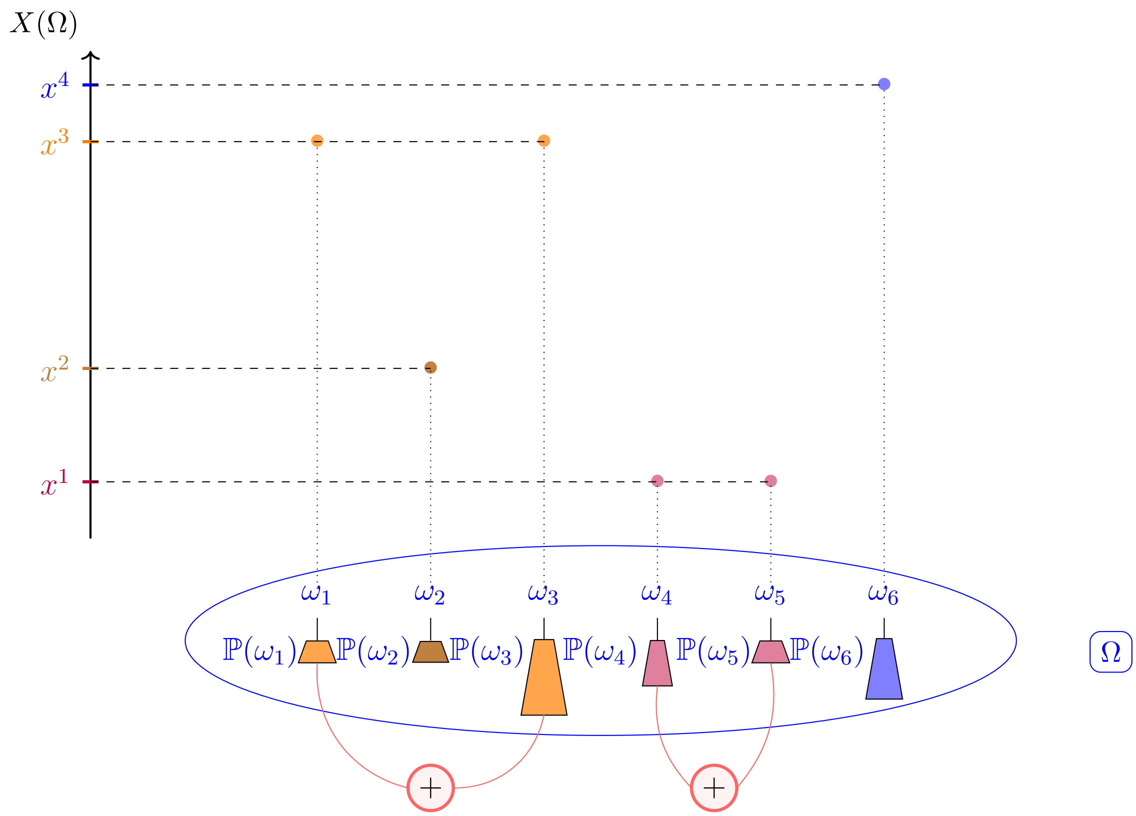

FIGURE 1 – Les dessous de la formule de transfert. Les probabilités des événements élémentaires de l'univers Ω sont ici figurées par de petits poids.

Partant de (2),

$$\begin{aligned}\mathrm{E}[X] = x^3\mathbb{P}(\omega_1) + x^2\mathbb{P}(\omega_2) + x^3\mathbb{P}(\omega_3)\\ + x^1\mathbb{P}(\omega_4) + x^1\mathbb{P}(\omega_5) + x^4\mathbb{P}(\omega_6)\end{aligned}$$

on débouche, en ré assemblant, sur

$$\begin{aligned}\mathrm{E}[X] = x^1\left[\mathbb{P}(\omega_4) + \mathbb{P}(\omega_5)\right] + x^2\mathbb{P}(\omega_2)\\ + x^3\left[\mathbb{P}(\omega_1) + \mathbb{P}(\omega_3)\right] + x^4\mathbb{P}(\omega_6)\end{aligned}$$

Soit exactement (1) :

$$\begin{aligned}\mathrm{E}[X] = x^1\mathbb{P}_X([X = x^1]) + x^2\mathbb{P}_X([X = x^2])\\ + x^3\mathbb{P}_X([X = x^3]) + x^4\mathbb{P}_X([X = x^4])\end{aligned}$$

Mais qui est Ω ?

Pour des raisons historiques, le lexique des probabilités trouve souvent un écho dans celui de l'analyse. Ainsi une *variable aléatoire* est-elle une *fonction* (mesurable) d'un espace vers un autre. L'usage a voulu noter cet espace « de départ » Ω et l'appeler *univers* ou *espace des possibles* – bien sûr, il n'a rien d'unique et dépend de la situation modélisée. Dans les cas simples (comme le lancer d'un dé) on y accède directement (ici $\Omega = \{1, 2, 3, 4, 5, 6\}$). Dans d'autres, plus complexes, il peut se faire plus cachotier. Il est alors déraisonnable (mais heureusement, pas forcément utile), de le caractériser. Par exemple, si l'on choisit pour Ω toutes les trajectoires possibles des gouttes de pluies et que l'on s'attache uniquement au nombre d'entre elles qui tombent dans un verre d'eau pendant une heure, c'est plutôt le point d'arrivée : le verre d'eau ou hors du verre d'eau, qui nous intéressera. Et pas Ω, du reste nuageux !

Références

[1] Marie-Line CHABANOL et Jean-Jacques RUCH. *Probabilités et statistiques*. Ellipses, 2016.

[2] Jacques ISTAS. *Probabilités et statistiques*. Ellipses, 1999.

[3] Jean-François LE GALL. *Intégration, probabilités et processus aléatoires*. Édition de l'École normale supérieure, 2006.

[4] Sylvie MÉLÉARD. *Aléatoire*. Les éditions de l'École polytechnique, 2014.

[5] Jean-Yves OUVRARD. *Probabilités 2*. Cassini, 2009.

[6] « Programme d'enseignement de spécialité de mathématiques de la classe de Première de la voie générale ». In : *Bulletin officiel spécial n°1 du 22 janvier 2019* (2019). URL : `https://www.education.gouv.fr/bo/19/Special1/MENE1901632A.htm`.

[7] « Programme de l'enseignement de spécialité de mathématiques de la classe de Terminale de la voie générale ». In : *Bulletin officiel spécial n°8 du 25 juillet 2019* (2019). URL : `https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm`.

[8] Olivier RIOUL. *Théorie des probabilités*. Hermès – Lavoisier, 2008.