Temps d'attente moyen d'une suite de pile ou face

Pierre Carriquiry(*)

À la suite de l'article de Jean-Paul Delahaye « Quand l'intuition a tout faux », paru dans le bulletin 504, je voudrais signaler une méthode qui permet de calculer le temps d'attente moyen d'une suite de pile ou face sans utiliser la notion d'espérance conditionnelle, et qui donne une formule générale qui permet de justifier des résultats donnés dans l'article.

On utilisera ici une formule donnant l'espérance d'une variable aléatoire X à valeurs entières positives :

$$\mathrm{E}(\mathrm{X})=\sum_{k=1}^{\infty} k\,\mathrm{P}(\mathrm{X}=k)=\sum_{k=0}^{\infty}\mathrm{P}(\mathrm{X}>k).$$

I Cas partiulier : temps d'attente moyen de PFPF

Soit X la variable aléatoire : « Temps d'attente de PFPF », c'est-à-dire le rang de l'expérience où on obtient pour la première fois PFPF. Les valeurs possibles de X sont les entiers supérieurs ou égaux à 4. Soit k un entier supérieur ou égal à 8, et A_k l'événement : « on obtient PFPF à la k^{e} épreuve. » (pas forcément pour la première fois: on obtient F à la k^{e}, P à la $(k-1)^{\mathrm{e}}$, F à la $(k-2)^{\mathrm{e}}$ et P à la $(k-3)^{\mathrm{e}}$). On a :

$$\begin{aligned}\mathrm{P}(\mathrm{A}_k)=\mathrm{P}(\mathrm{X}=k)\times\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k)+\mathrm{P}(\mathrm{X}=k-1)\times\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-1)\\+\cdots+\mathrm{P}(\mathrm{X}=k-4)\times\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-4)\end{aligned}\qquad(1)$$

$\mathrm{P}(\mathrm{A}_k)=\left(\frac{1}{2}\right)^4$; $\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k)=1$; $\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-1)=0$ car si la séquence PFPF apparaît pour la première fois au rang $(k-1)$ elle ne peut pas apparaître ensuite au rang k.

$\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-2)=\left(\frac{1}{2}\right)^2$ car si la séquence PFPF apparaît pour la première fois au rang $(k-2)$, elle apparaîtra aussi au rang k si les 2 épreuves suivantes donnent PF.

De même, $\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-3)=0$ et pour tout $i\in\{4,\ldots,(k-4)\}$,

$$\mathrm{P}(\mathrm{A}_k/\mathrm{X}=i)=\left(\frac{1}{2}\right)^4.$$

(*) carriquiry.pierre@neuf.fr

L'égalité (1) donne alors :

$$\left(\frac{1}{2}\right)^4 = \mathrm{P}(\mathrm{X}=k)+\left(\frac{1}{2}\right)^2 \mathrm{P}(\mathrm{X}=k-2)+\left(\frac{1}{2}\right)^4 \mathrm{P}(\mathrm{X}\leq k-4)$$

On en tire :

$$\left(\frac{1}{2}\right)^4 \mathrm{P}(\mathrm{X}>k-4)=\mathrm{P}(\mathrm{X}=k)+\left(\frac{1}{2}\right)^2 \mathrm{P}(\mathrm{X}=k-2)$$

D'où :

$$\mathrm{P}(\mathrm{X}>k-4)=2^4\,\mathrm{P}(\mathrm{X}=k)+2^2\,\mathrm{P}(\mathrm{X}=k-2)$$

et on vérifie que cette égalité est encore valable pour k = 7, 6, 5 et 4. Par exemple pour k = 7, le raisonnement qui conduit à l'égalité (1) donne :

$$\left(\frac{1}{2}\right)^4 = \mathrm{P}(\mathrm{X}=7)+\left(\frac{1}{2}\right)^2 \mathrm{P}(\mathrm{X}=7-2)$$

d'où

$$1=2^4\,\mathrm{P}(\mathrm{X}=7)+2^2\,\mathrm{P}(\mathrm{X}=5),$$

et $\mathrm{P}(\mathrm{X}>7-4)=1.$

On a alors :

$$\mathrm{E}(\mathrm{X})=\sum_{k=4}^{+\infty}\mathrm{P}(\mathrm{X}>k-4)=2^4\sum_{k=4}^{+\infty}\mathrm{P}(\mathrm{X}=k)+2^2\sum_{k=4}^{+\infty}\mathrm{P}(\mathrm{X}=k-2)=2^4+2^2=20.$$

II Cas général

Soit S = $(u_1,u_2,\cdots,u_n)$ une suite de pile ou face, (u_i = P ou F), X la variable aléatoire « Temps d'attente de S », k un entier supérieur ou égal à $2n$, A_k l'événement : « $(u_1,u_2,\cdots,u_n)$ apparaît à la k^{e} épreuve".

On a :

$$\begin{aligned}\mathrm{P}(\mathrm{A}_k)=\mathrm{P}(\mathrm{X}=k)+\mathrm{P}(\mathrm{X}=k-1)\times\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-1)\\ +\cdots+\mathrm{P}(\mathrm{X}=n)\times\mathrm{P}(\mathrm{A}_k/\mathrm{X}=n)\end{aligned}\qquad(2)$$

$\mathrm{P}(\mathrm{A}_k)=\left(\frac{1}{2}\right)^n$. Pour tout $i\in\{0, 1, \ldots, n-1)\}$, $\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-i)=\left(\frac{1}{2}\right)^i$ si la suite S peut apparaître aux rangs $k-i$ et k, et $\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-i)=0$ sinon. Pour tout $i\in\{n, n+1, \ldots, k-n)\}$, $\mathrm{P}(\mathrm{A}_k/\mathrm{X}=k-i)=\left(\frac{1}{2}\right)^n$ ..

On définit l'ensemble d'indices $\mathrm{I}=\{0\leq i<n, \mathrm{P}(\mathrm{A}_k / \mathrm{X}=k-i)\neq 0\}$..L'égalité (2) donne alors :

$$\left(\frac{1}{2}\right)^n=\sum_{i\in\mathrm{I}}\left(\frac{1}{2}\right)^i \mathrm{P}(\mathrm{X}=k-i)+\left(\frac{1}{2}\right)^n \mathrm{P}(\mathrm{X}\leq k-n),$$

d'où on tire :

$$\mathrm{P}(\mathrm{X}>k-n)=\sum_{i\in\mathrm{I}}2^{n-i}\,\mathrm{P}(\mathrm{X}=k-i)$$

et cette égalité est vérifiée pour $n \leq k < 2n$ car dans ce cas (2) se réduit à :

$$\left(\frac{1}{2}\right)^n=\sum_{i\in\mathrm{I}}\left(\frac{1}{2}\right)^i \mathrm{P}(\mathrm{X}=k-i)$$

et $\mathrm{P}(\mathrm{X}>k-n)=1$

On a alors

$$\mathrm{E}(\mathrm{X})=\sum_{k=n}^{+\infty}\mathrm{P}(\mathrm{X}>k-n)=\sum_{i\in\mathrm{I}}2^{n-i}\sum_{k=n}^{+\infty}\mathrm{P}(\mathrm{X}=k-i)=\sum_{i\in\mathrm{I}}2^{n-i}$$

car $\sum_{k=n}^{+\infty}\mathrm{P}(\mathrm{X}=k-i)=1$ pour $i < n$.

Pour simplifier un peu plus, on fait le changement d'indice $n - i = j$ et on obtient :

$$\mathrm{E}(\mathrm{X})=\sum_{j\in\mathrm{J}}2^j$$

où $\mathrm{J}=\{1\leq j\leq n, \mathrm{P}(\mathrm{A}_{2n} / \mathrm{X}=n+j)\neq 0\}$. est l'ensemble des indices j tels que la suite S peut apparaître aux rangs $2n$ et $n + j$.

Par exemple, si S = (FF), J = {1, 2} et E(X) = $2^2 + 2 = 6$.

Si S = (PF), J = {2} et E(X) = $2^2 = 4$.

Si S = (FPPF), J = {1;4} et E(X) = $2^4 + 2^1 = 18$.

Si S = (FF...F) (n temes), J = {1, 2, …, n} et E(X) = $2^n + 2^{n-1} + \ldots + 2 = 2^{n+1} - 2$.

Si S = (PF...F) (n termes), J = {n} et E(X) = 2^n.

On a toujours $n \in$ J, donc le temps d'attente moyen d'une suite de n termes est compris entre 2^n et $2^{n+1} - 2$.

III Cas d'une pièce non équilibrée

En utilisant la méthode précédente, on trouve que le temps moyen d'attente de la suite $(u_1, u_2, \cdots, u_n)$ est

$$\mathrm{E}(\mathrm{X})=\sum_{j\in\mathrm{J}}\frac{1}{\mathrm{P}(u_1)\mathrm{P}(u_2)\cdots\mathrm{P}(u_j)}$$

où $\mathrm{J} = \left\{ j \in \{1,2,\cdots,n\}, \forall i \in \{0,1,\cdots,j-1\} u_{n-i} = u_{j-i} \right\}$ est l'ensemble des indices j tels que la suite formée par les j premiers termes soit égale à la suite formée par les j derniers.

Par exemple, si on note p la probabilité d'obtenir pile, le temps d'attente moyen de FF est $\dfrac{1}{(1-p)^2} + \dfrac{1}{1-p}$, celui de PF est $\dfrac{1}{p(1-p)}$; on trouve qu'ils sont égaux pour $p = \dfrac{3-\sqrt{5}}{2}$ et le temps d'attente moyen est alors $\sqrt{5}+2 = 4{,}236\ldots$. On n'est pas étonné de ne pas trouver 5 (qui est la moyenne des temps d'attente moyens de FF et PF pour une pièce équilibrée) car « l'intuition a tout faux », mais on est un peu surpris de remarquer que si $\varphi = \dfrac{\sqrt{5}+1}{2}$, les temps moyens d'attente de FF et PF sont égaux quand la probabilité de face est , $\dfrac{1}{\varphi}$ celle de pile $\dfrac{1}{\varphi^2}$ et le temps d'attente moyen est alors φ^3. Mais que vient faire le nombre d'or dans ce calcul de probabilité ? Un hasard, probablement, à moins qu'un lecteur ne nous explique (intuitivement) pourquoi le nombre d'or doit figurer de manière aussi simple dans ces résultats, ce qui réhabiliterait un peu l'intuition.