

Énoncé du problème

On lance une pièce parfaitement équilibrée jusqu'à obtenir "Face" deux fois de suite. Combien de lancers en moyenne va-t-on effectuer ?

Solution

Selon l'habitude, dans cette expérience aléatoire, nous prendrons pour modèle de l'ensemble des événements élémentaires : $\Omega = \{0;1\}^{\mathbb{N}^*}$. Nous noterons les variables aléatoires dites de Bernoulli : $(X_i)_{i \in \mathbb{N}^*}$, telles que : $X_i(\omega) = \omega(i) = \omega_i$ pour tout $i \in \mathbb{N}^*$. L'algèbre des événements envisagée, est la plus petite algèbre rendant mesurable toutes les X_i . La probabilité sur cette algèbre d'événements suppose que toutes les X_i sont deux à deux indépendantes, avec pour fonction de répartition :

ω_i	1	0
$P(X_i = \omega_i)$	$p = \frac{1}{2}$	$q = \frac{1}{2}$

Un terme de rang i égal à 1 symbolisera l'apparition de «face» lors du i^{e} tirage. Considérons les variables aléatoires Y_n pour $n \in \mathbb{N}$ et $n \geq 2$, qui comptabilisent lors des n premiers tirages, le nombre d'apparitions de «face», alors que le tirage précédent avait déjà fait apparaître «face», plus précisément :

$$Y_n = \sum_{i=2}^n X_i X_{i-1}$$

Par exemple si $\omega = (1,0,1,1,0,1,1,0,\dots)$: $Y_3(\omega) = 0$; $Y_4(\omega) = Y_5(\omega) = Y_6(\omega) = 1$; $Y_7(\omega) = 2$ et $Y_8(\omega) = Y_9(\omega) = 3$. Il est important de remarquer pour les raisonnements qui vont suivre, que pour tout ω , $Y_n(\omega)$ est une fonction croissante de n . Le problème posé consiste alors à calculer l'espérance mathématique de la variable aléatoire Z à valeurs entières, telle que $Z(\omega) = \inf\{n \in \mathbb{N} / Y_n(\omega) = 1\}$, elle indique le rang à partir duquel on a rencontré la première apparition de 2 «faces» consécutifs. Dans l'exemple précédent : $Z(\omega) = 4$. On a de manière évidente :

- $\{Z \leq 1\}$ impossible,
- $\{Z = 2\}$ réalisé lorsque les deux premiers tirages ont donné «face», donc : $P(Z = 2) = p^2$.
- $\{Z = 3\}$ est réalisé pour les tirages commençant par $(0;1;1\dots)$, donc : $P(Z = 3) = p^2 q$.

Pour $Z = n > 3$ un calcul plus élaboré est nécessaire¹ :

$$P(Z = n) = P(\{Y_{n-1} = 0\} \cap \{Y_n = 1\}) = P(Y_{n-1} = 0) \times P(Y_n = 1 | Y_{n-1} = 0)$$

Sachant : $Y_{n-1} = 0$, on a : $Y_n = X_{n-1} X_n$, donc : $Y_n = 1 \Leftrightarrow X_n = 1 \text{ et } X_{n-1} = 1$, on en déduit :

$$P(Z = n) = P(Y_{n-1} = 0) \times p^2$$

D'après le théorème de Bayes :

$$P(Y_{n-1} = 0) = q.P(Y_{n-1} = 0 | X_{n-1} = 0) + p.P(Y_{n-1} = 0 | X_{n-1} = 1)$$

Sachant : $X_{n-1} = 0$, on a : $Y_{n-1} = Y_{n-2}$, donc : $Y_{n-1} = 0 \Leftrightarrow Y_{n-2} = 0$, on en déduit :

$$P(Y_{n-1} = 0 | X_{n-1} = 0) = P(Y_{n-2} = 0)$$

Sachant : $X_{n-1} = 1$, on a : $Y_{n-1} = Y_{n-3} + X_{n-3} \times X_{n-2} + X_{n-2}$, donc : $Y_{n-1} = 0 \Leftrightarrow (X_{n-2} = 0 \text{ et } Y_{n-3} = 0)$

X_{n-2} est indépendante de la variable Y_{n-3} qui s'exprime à partir des $(X_i)_{i \leq n-3}$, on en déduit :

$$P(Y_{n-1} = 0 | X_{n-1} = 1) = q.P(Y_{n-3} = 0)$$

On a donc les formules de récurrence : $P(Y_{n-1} = 0) = q.P(Y_{n-2} = 0) + p.q.P(Y_{n-3} = 0)$

$$P(Z = n) = q.P(Z = n-1) + p.q.P(Z = n-2)$$

On a donc : $E(Z) = \sum_{n=2}^{+\infty} n.z_n$ où la suite $(z_n)_{n \geq 2}$ est définie par récurrence linéaire :

$$\begin{cases} z_2 = p^2 \\ z_3 = q.p^2 \\ z_n = q.z_{n-1} + p.q.z_{n-2} \end{cases}$$

1. On note $P(B|A) = \frac{P(B \cap A)}{P(A)}$, la probabilité de B sachant A

La théorie des suites récurrentes linéaires permet un calcul explicite de z_n que je remets à plus tard. Lorsque $p = q = \frac{1}{2}$, la suite (z_n) s'exprimera plus simplement à l'aide de la classique suite récurrente de Fibonacci² dont une expression explicite en fonction du nombre d'or est connue. Cette suite $(f_n)_{n \in \mathbb{N}}$ est définie par :

$$\begin{cases} f_0 &= & 0 \\ f_1 &= & 1 \\ f_n &= & f_{n-1} + f_{n-2} \end{cases}.$$

On montre alors par récurrence que lorsque $p = \frac{1}{2}$ on a : $z_n = \frac{f_{n-1}}{2^n}$ pour $n \geq 2$;

- nous avons déjà calculer : $z_2 = p^2 = \frac{1}{4} = \frac{f_1}{2^2}$ et $z_3 = p^2 q = \frac{1}{8} = \frac{f_2}{2^3}$.
- Pour $n > 3$, on a : $z_n = q \cdot z_{n-1} + p q \cdot z_{n-2} = \frac{1}{2} \times \frac{f_{n-2}}{2^{n-1}} + \frac{1}{4} \times \frac{f_{n-3}}{2^{n-2}} = \frac{f_{n-1}}{2^n}$

Nous avons l'expression explicite $f_n = \frac{\varphi_1^n - \varphi_2^n}{\sqrt{5}}$ où φ_1 est le nombre d'or $\frac{1+\sqrt{5}}{2}$ et φ_2 l'opposé de son inverse : $\frac{1-\sqrt{5}}{2}$, cela nous permet d'exprimer l'espérance de Z par la série :

$$E(Z) = \frac{1}{2\sqrt{5}} \sum_{n=2}^{+\infty} (n (x_1^{n-1} - x_2^{n-1})) \quad \text{où} \quad x_1 = \frac{\varphi_1}{2} = \frac{1+\sqrt{5}}{4} \quad \text{et} \quad x_2 = \frac{\varphi_2}{2} = \frac{1-\sqrt{5}}{4}$$

Or : $\sum_{n=2}^{+\infty} n x^{n-1} = f'(x)$ où $f(x)$ est la somme de la série entière $\sum_{n=2}^{+\infty} x^n = \frac{x^2}{1-x}$, dont le rayon de convergence est 1 qui majore

strictement $|x_1|$ et $|x_2|$; on a donc : $f'(x) = \frac{2x - x^2}{(1-x)^2}$. Pour calculer $f'(x_1)$ et $f'(x_2)$, posons $x = \frac{1+\varepsilon\sqrt{5}}{4}$ où $\varepsilon \in \{1; -1\}$, on a :

$$2x - x^2 = \frac{8 + 8\varepsilon\sqrt{5} - 6 - 2\varepsilon\sqrt{5}}{16} = \frac{2 + 6\varepsilon\sqrt{5}}{16} \quad \text{et} \quad (1-x)^2 = \frac{(3 - \varepsilon\sqrt{5})^2}{16} = \frac{14 - 6\varepsilon\sqrt{5}}{16}$$

$$\text{On en déduit : } f'(x_1) = \frac{2 + 6\sqrt{5}}{14 - 6\sqrt{5}} \quad \text{et} \quad f'(x_2) = \frac{2 - 6\sqrt{5}}{14 + 6\sqrt{5}}$$

$$\text{Donc : } E(Z) = \frac{1}{2\sqrt{5}} \times \frac{(2 + 6\sqrt{5})(14 + 6\sqrt{5}) - (2 - 6\sqrt{5})(14 - 6\sqrt{5})}{14^2 - 36 \times 5}$$

$$E(Z) = \frac{1}{2\sqrt{5}} \times \frac{2(2 \times 6\sqrt{5} + 6 \times 14\sqrt{5})}{16} = \frac{1}{2\sqrt{5}} \times \frac{2 \times 96\sqrt{5}}{16}$$

$$\text{On obtient finalement : } E(Z) = 6$$

2. http://fr.wikipedia.org/wiki/Suite_de_Fibonacci

Ci-joint un classeur Excel conçue en 2000-2001 pour expérimenter divers comportements du jeu de pile ou face. Ce classeur comporte 1000 lignes, qui permettent chacune d'observer une série de 400 tirages se déroulant progressivement. Je l'ai légèrement modifiée pour y inclure une colonne où figure le rang du premier motif que l'on veut voir apparaître, '11' par exemple pour 2 faces consécutifs. Il faut taper ce motif entre cotes, pour que '11', '01' ou '000' soient interprétés comme chaînes de caractères et non comme nombres.

Open Office n'existait pas en 2000, et je n'ai jamais eu le courage d'adapter les macros Excel à OpenOffice. Cette feuille de calcul ne fonctionne donc que sous Excel, en autorisant l'exécution des macros.

L'expérience est très significative, car en appuyant sur les boutons associés aux macros «remise à zéro» et «continuer les tirages», on relance en parallèle 1000 jeux, en effectuant d'abord 25 tirages, puis 50, 75 et ainsi de suite jusqu'à 400 tirages, par tranche de 25. On remarque alors pour les 25 premiers tirages une moyenne sur les 100 échantillons, toujours inférieure à 6, il faut continuer les tirages jusqu'à 100 pour voir la moyenne se stabiliser à un nombre presque toujours proche de 6 à 0,1 près. Après 200 tirages la moyenne est stabilisée, car il est très improbable qu'il y ait encore une série parmi 1000, sans «double faces» intervenu avant 200 tirages.

