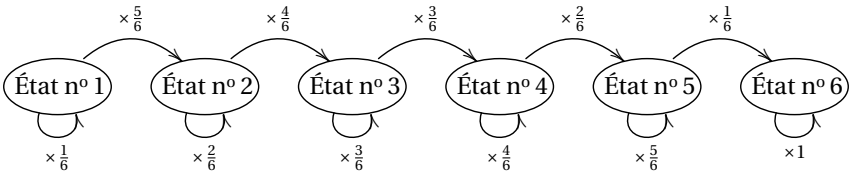
Énoncé du problème

En moyenne, combien de fois faut-il lancer un dé parfaitement équilibré, pour finir par obtenir au moins une fois les 6 faces possibles ?

Solution

Soit $\Omega = \{1;2;3;4;5;6\}^{\mathbb{N}^*}$ l'ensemble des suites $(\omega_1, \omega_2, \ldots)$, modélisant chacune, une suite infinie de lancer d'un dé, où $\omega_i \in \{1;2;3;4;5;6\}$ désigne la face obtenue lors du lancer i. La mesure ou probabilité définie sur Ω est celle rendant mesurables et indépendantes les variables aléatoires $X_i : \Omega \longrightarrow \mathbb{N}$ telles que $X_i(\omega) = \omega_i$ Le problème posé consiste à calculer l'espérance de la variable aléatoire $Y : \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{N} \\ \omega & \longmapsto & \sup\limits_{1 \leqslant i \leqslant 6} \left[\inf\left(\omega^{-1}(i)\right)\right] \end{array}$.

$Y(\omega)$ ainsi défini est le plus petit entier à partir duquel est apparue au moins une fois, chacune des 6 éventualités [1]. Deux méthodes sont proposées pour calculer l'expérance de Y, chacune fait appel à la description suivante du processus aléatoire à partir d'un graphe probabiliste.



À l'instant n, le processus est dit dans l'état $i \in \{1;2;3;4;5;6\}$ lorsque seulement i faces sont apparues au cours des n premiers lancers, les poids indiqués sur chaque flèche du graphe indique les probabilités de rester dans l'état i ou de passer de l'état i à l'état $i+1$. C'est l'indépendance des variables X_i découlant de l'hypohèse qu'aucun lancer peut influencer un autre lancer, qui permet d'évaluer ces probabilités. Dans la première méthode nous déterminerons la loi de répartition de la variable aléatoire Y, illustrée par la simulation sous excel jointe à ce pdf. Nous verrons une deuxième méthode de calcul qui permet de calculer son espérance sans connaitre sa répartition. Cette méthode plus courte ne permet pas toutefois de répondre à d'autres questions telles que

- Calcul de l'écart-typeet et d'intervalles de fluctuation.
- À partir de quel lancer est-il le plus probable que soient apparues pour la première fois les 6 faces ?

1^e^méthode : théorie des chaines de Markov

Ce processus aléatoire est une chaine de Markov à 6 états, dont l'état n° 6 est l'état stable vers lequel converge le processus. Nous avons besoin pour l'étudier de définir par récurrence, les 6 suites ci-dessous, qui indiquent respectivement la probabilité d'être à un instant n, dans l'état 1 , 2 , 3 , 4 , 5 ou 6. Elles sont la formulation mathématique du graphe ci-dessus, après avoir appliqué le théorème des probabilités totales.

- $u_1 = 1$ et $u_{n+1} = \frac{1}{6}u_n$
- $d_1 = 0$ et $d_{n+1} = \frac{2}{6}d_n + \frac{5}{6}u_n$
- $t_1 = 0$ et $t_{n+1} = \frac{3}{6}t_n + \frac{4}{6}d_n$
- $q_1 = 0$ et $q_{n+1} = \frac{4}{6}q_n + \frac{3}{6}t_n$
- $c_1 = 0$ et $c_{n+1} = \frac{5}{6}c_n + \frac{2}{6}q_n$
- $s_1 = 0$ et $s_{n+1} = s_n + \frac{1}{6}c_n$

La probabilité que ce processus passe de l'état n° 5 à l'état stable n° 6 juste après le lancer n° $(n+1)$ est : $s_{n+1} - s_n = \dfrac{1}{6}c_n$, nous avons simplement besoin de l'expression explicite de c_n, pour calculer l'espérance de Y qui est égale à $\dfrac{1}{6}\sum\limits_{n=5}^{+\infty}(n+1)c_n$, ce que nous pouvons faire à l'aide du calcul matriciel suivant.

Les 6 suites donnant les probabilités d'être dans l'un de ces 6 états, peuvent être définies par récurrence à l'aide de ce

1. Cette façon de définir X permet une simulation sous tableur disponible à l'URL suivante : http://lyc-marguerite-valois-math.fr/pb15/6faces.ods

produit matriciel suivant :

$$\begin{pmatrix} u_{n+1} \\ d_{n+1} \\ t_{n+1} \\ q_{n+1} \\ c_{n+1} \\ s_{n+1} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 & 0 & 0 \\ 0 & 4 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 4 & 0 & 0 \\ 0 & 0 & 0 & 2 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{pmatrix}\begin{pmatrix} u_n \\ d_n \\ t_n \\ q_n \\ c_n \\ s_n \end{pmatrix} \text{ On en déduit : } \begin{pmatrix} u_{n+1} \\ d_{n+1} \\ t_{n+1} \\ q_{n+1} \\ c_{n+1} \\ s_{n+1} \end{pmatrix} = \frac{1}{6^n}\mathscr{M}^n \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

où $\mathscr{M}$ est la matrice 6×6 ci-dessus, telle que $\frac{1}{6}\mathscr{M}$ soit la matrice de transition du processus aléatoire. Les valeurs propres de cette matrice triangulaire sont toutes distinctes, ce sont les 6 coefficients de sa diagonale. Les colonnes de la matrice de passage $\mathscr{P}$ permettant de diagonaliser $\mathscr{M}$, est constituée des coordonnées des 6 vecteurs propres dans la base canonique de $\mathbb{R}^6$. On les obtient aisément en résolvant les 6 systèmes $\mathscr{M}\mathscr{X} = n\mathscr{X}$, en l'inconnue $\mathscr{X}$ pour $n \in [\![1;6]\!]$

$$\mathscr{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 & 0 \\ 10 & -4 & 1 & 0 & 0 & 0 \\ -10 & 6 & -3 & 1 & 0 & 0 \\ 5 & -4 & 3 & -2 & 1 & 0 \\ -1 & 1 & -1 & 1 & -1 & 1 \end{pmatrix} \text{ ; son inverse est : } \mathscr{P}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 & 0 & 0 \\ 10 & 4 & 1 & 0 & 0 & 0 \\ 10 & 6 & 3 & 1 & 0 & 0 \\ 5 & 4 & 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Seule la colonne n° 1 de $\mathscr{P}^{-1}$ suffit pour aboutir assez simplement, à l'expression explicite de nos six suites, résultant du produit maticiel suivant :

$$\begin{pmatrix} u_{n+1} \\ d_{n+1} \\ t_{n+1} \\ q_{n+1} \\ c_{n+1} \\ s_{n+1} \end{pmatrix} = \frac{1}{6^n}\mathscr{P}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2^n & 0 & 0 & 0 & 0 \\ 0 & 0 & 3^n & 0 & 0 & 0 \\ 0 & 0 & 0 & 4^n & 0 & 0 \\ 0 & 0 & 0 & 0 & 5^n & 0 \\ 0 & 0 & 0 & 0 & 0 & 6^n \end{pmatrix}\mathscr{P}^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{6^n}\mathscr{P}\begin{pmatrix} 1 \\ 5\times 2^n \\ 10\times 3^n \\ 10\times 4^n \\ 5\times 5^n \\ 6^n \end{pmatrix}$$

Nous avons simplement besoin de l'expression explicite de c_n. D'après le produit matriciel ci-dessus

$$c_{n+1} = \frac{1}{6^n}(5 - 20\times 2^n + 30\times 3^n - 20\times 4^n + 5\times 5^n)$$

$$\text{On a donc : } \frac{1}{6}c_n = \frac{1}{6^n}(5 - 10\times 2^n + 10\times 3^n - 5\times 4^n + 5^n)$$

On sait et on peut le vérifier sur cette expression explicite que c_n est nul pour $n \in \{1;2;3;4\}$. Pour simplifier le calcul de $E(Y)$, on peut créer le terme fictif $\frac{1}{6}c_0 = 1$ à soustraire, dans une sommation faite à partir de l'indice 0 :

$$E(Y) = \sum_{n=0}^{+\infty}\frac{(n+1)c_n}{6} - 1 = \sum_{n=0}^{+\infty}(n+1)\left(\frac{5}{6^n} - \frac{10}{3^n} + \frac{10}{2^n} - \frac{5\times 2^n}{3^n} + \frac{5^n}{6^n}\right) - 1$$

Or $f(x) = \sum_{n=0}^{+\infty}(n+1)x^n$ est une fonction analytique de rayon de convergence 1, résultant de la dérivée de $F(x) = \sum_{n=0}^{+\infty}x^n = \frac{1}{1-x}$, on a donc $f(x) = \frac{1}{(1-x)^2}$ qui nous permet le calcul des 5 rationnels suivants :

$$f\left(\frac{1}{6}\right) = \frac{36}{25} \quad ; \quad f\left(\frac{1}{3}\right) = \frac{9}{4} \quad ; \quad f\left(\frac{1}{2}\right) = 4 \quad ; \quad f\left(\frac{2}{3}\right) = 9 \quad ; \quad f\left(\frac{5}{6}\right) = 36$$

$$\text{Donc : } E(Y) = 5f\left(\frac{1}{6}\right) - 10f\left(\frac{1}{3}\right) + 10f\left(\frac{1}{2}\right) - 5f\left(\frac{2}{3}\right) + f\left(\frac{5}{6}\right) - 1 = \frac{36}{5} - \frac{45}{2} + 40 - 45 + 36 - 1$$

Finalement on obtient que le nombre moyen de lancers de dé pour finir par obtenir au moins une fois chacune des 6 faces est le nombre décimal **14,7**.

2^e^méthode accessible aux lycéens connaisssant les variables aléatoires géométriques.

Cette méthode permet de calculer l'espérance de Y, sans être obligé d'expliciter sa loi de répartition. considérons les variables aléatoires Z_i, pour $i \in \{1;2;3;4;5\}$, telles que Z_i comptabilise le nombre de lancers entre les instants où l'on passe de l'état i à l'état $i+1$. Le premier lancer aboutit obligatoirement à l'état 1, Le nombre de lancers pour parvenir à l'état n° 6 est donc : $Y = 1 + Z_1 + Z_2 + Z_3 + Z_4 + Z_5$. On a donc $E(Y) = 1 + E(Z_1) + E(Z_2) + E(Z_3) + E(Z_4) + E(Z_5)$. Le fait que les variables X_i soient indépendantes permet d'affirmer : $P(Z_i = k) = \dfrac{6-i}{6}\left(\dfrac{i}{6}\right)^{k-1}$; car il faut qu'au k^e lancer après l'instant où le processus est passé dans l'état i, le dé tombe sur une des $(6-i)$ faces non encore obtenues, alors que les $(k-1)$ lancers précédents sont tous tombés sur une des i faces déjà intervenues. On a donc :

$$E(Z_i) = \frac{6-i}{6}\sum_{k=1}^{+\infty} k\left(\frac{i}{6}\right)^{k-1}$$

La somme des cette série peut être obtenue par dérivation de la fonction analytique de rayon de convergence 1 :

$$F(z) = \sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$$

$$\text{On a donc : } F'(z) = \sum_{k=1}^{\infty} kz^{k-1} = \frac{1}{(1-z)^2}$$

$$\text{D'où :} \qquad \frac{6-i}{6}\sum_{k=1}^{+\infty} k\left(\frac{i}{6}\right)^{k-1} = \frac{6-i}{6}F'\left(\frac{i}{6}\right) = \frac{6-i}{6} \times \frac{1}{\left(1-\frac{i}{6}\right)^2} = \frac{6}{6-i}$$

$$\text{Donc :} \qquad E(Z) = 1 + \frac{6}{5} + \frac{3}{2} + 2 + 3 + 6 = 14,7$$

Remarque : on ne peut pas calculer la variance de Y par addition, Car les Z_i ne sont pas indépendantes, et dans ce cas là, la variance n'est pas additive.