DM N$^{\text{o}}$ 15 : Un problème de probabilités

Lucas Tabary

Énoncé

On effectue une série de tirages avec remises dans une urne contenant n boules numérotées de 1 à n. On note X_k la variable aléatoire égale à la valeur de la boule tirée au k-ème tirage ($k > 0$). On note alors S_k la variable aléatoire comptabilisant la somme des valeurs obtenues jusqu'au k-ème tirage inclus. On a :

$$S_k = \sum_{i=1}^{k} X_i$$

On considère enfin la variable aléatoire T_n égale au nombre de tirages nécessaires pour que, pour la première fois, la somme des boules obtenues soit supérieure ou égale à n. L'étude se fera dans un espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

Partie A

1. Déterminons la loi de X_k. On remarque préalablement que tous les tirages sont réalisés de façon identique et indépendante. On peut dès lors considérer que chaque valeur a la même probabilité d'être désignée parmi n boules. Donc :

$$\forall k \in [\![1, n]\!],\ X_k \hookrightarrow \mathcal{U}_{[\![1,\,n]\!]} \Longrightarrow \mathbb{P}(X_k = k) = \frac{1}{n} \text{ et } \mathbb{E}(X_k) = \frac{n+1}{2}$$

2. (a) Déterminons la valeur de $T_n(\Omega)$. Dans le meilleur cas, on obtient la boule n au premier tirage, et alors $T_n = 1$. Dans le pire des cas on progresse au rythme de 1 par tirage, on obtiendra donc n au bout de n tirages, d'où $T_n = n$. Pour obtenir un nombre entre ces deux bornes, qu'on notera $k \in]\!]1, n[\![$, on peut tirer $k-1$ fois la boule 1 puis une fois la boule $n - k + 1 \in [\![1, n]\!]$, ce qui donnera bien une valeur totale de n pour k tirages : toutes les valeurs sont atteignables et on a : $T_n(\Omega) = [\![1, n]\!]$.

   (b) Le seul moyen d'obtenir n en 1 coup est de tirer la boule n, au premier coup *(sic)*. Soit, par équiprobabilité : $\mathbb{P}(T_n = 1) = \frac{1}{n}$.

   (c) On réalise n tirages pour obtenir une valeur totale supérieure ou égale à n. Au $n-1$-ème tirage, si une valeur est supérieure à 1, on a alors

$$\sum_{k=1}^{n-1} X_k > \sum_{k=1}^{n-1} 1 = n - 1 \Longrightarrow \sum_{k=1}^{n-1} X_k \geqslant n$$

Ce qui est absurde car on aurait alors $T_n = n-1$. Le dernier n'a pas d'importance car dans tous les cas on obtient une valeur supérieure ou égale à n. D'où :

$$\mathbb{P}(T_n = n) = \left(\frac{1}{n}\right)^{n-1} \cdot 1$$

3. Pour T_2, on exploite directement les résultats des questions précédentes : on a alors :
   — $\mathbb{P}(T_2 = 1) = \frac{1}{2}$;
   — $\mathbb{P}(T_2 = 2) = \left(\frac{1}{2}\right)^{2-1} = \frac{1}{2}$.
4. Pour T_3, on utilise encore les formules :
   — $\mathbb{P}(T_3 = 1) = \frac{1}{3}$;
   — $\mathbb{P}(T_3 = 3) = \left(\frac{1}{3}\right)^{3-1} = \frac{1}{9}$;
   — En utilisant le système complet d'évènements associé à T_3,

$$\mathbb{P}(T_3 = 2) = 1 - [\mathbb{P}(T_3 = 1) + \mathbb{P}(T_3 = 3)] = 1 - \frac{4}{9} = \frac{5}{9}$$

   D'où $\mathbb{E}(T_3) = 1 \cdot \frac{1}{3} + 3 \cdot \frac{1}{9} + 2 \cdot \frac{5}{9} = \frac{2}{3} + \frac{10}{9} = \frac{16}{9}$.

Partie B

5. On reconsidère les pires et meilleurs cas décrit dans la question 2a. On a alors :

$$k = \sum_{i=1}^{k} 1 \leqslant S_k = \sum_{i=1}^{k} X_i \leqslant \sum_{i=1}^{k} n = kn$$

   Toutes les valeurs étant atteignables, on a : $S_k(\Omega) = [\![k; kn]\!]$.
6. (a) Par définition de S_k sur l'intervalle considéré :

$$S_{k+1} = \sum_{i=1}^{k+1} X_i = X_{k+1} + \sum_{i=1}^{k} X_i = X_{k+1} + S_k$$

   (b) On considère le système complet d'évènements associé à S_k, soit $\{(S_k = j), j \in [\![k; kn]\!]\}$, on remarque que $X_{k+1} = S_{k+1} - S_k$, on utilise alors la loi des probabilités totales :

$$\mathbb{P}(S_{k+1} = i) = \sum_{j=k}^{kn} \underbrace{\mathbb{P}_{(S_k=j)}(S_{k+1} = i)}_{\mathbb{P}(X_{k+1}=i-j)} \mathbb{P}(S_k = j)$$

   Or $\mathbb{P}(X_{k+1} = i-j) = \begin{cases} \dfrac{1}{n} \text{ si } i-j \in [\![1, n]\!] \\ 0 \text{ sinon} \end{cases}$. Les cas non nuls se présentent pour $1-n \leqslant j \leqslant i-1$.
   On remplace donc dans l'expression précédente et alors :

$$\mathbb{P}(S_{k+1} = i) = \sum_{j=k}^{i-1} \underbrace{\mathbb{P}(X_{k+1} = i-j)}_{\frac{1}{n}} \mathbb{P}(S_k = j) + \sum_{j=i}^{kn} \underbrace{\mathbb{P}(X_{k+1} = i-j)}_{0} \mathbb{P}(S_k = j) = \frac{1}{n} \sum_{j=k}^{i-1} \mathbb{P}(S_k = j)$$

7. (a) On rappelle directement la relation de Pascal, pour $0 < k < n$:

$$\binom{j}{k} = \binom{j-1}{k-1} + \binom{j-1}{k} \Longrightarrow \binom{j-1}{k-1} = \binom{j}{k} - \binom{j-1}{k}$$

(b) On somme les deux termes de l'égalité précédente de $j = k$ à $i-1$ (avec nécessairement $i - 1 \geqslant k$) :

$$\sum_{j=k}^{i-1} \binom{j-1}{k-1} = \sum_{j=k}^{i-1} \binom{j}{k} - \sum_{j=k}^{i-1} \binom{j-1}{k} = \sum_{j=k+1}^{i} \binom{j-1}{k} - \sum_{j=k}^{i-1} \binom{j-1}{k} = \binom{i-1}{k} - \underbrace{\binom{k-1}{k}}_{0}$$

On obtient alors la formule demandée.

(c) On considère $\mathcal{H}_k$, qu'on se propose de démontrer par récurrence forte, pour k un entier dans $[\![1, n]\!]$.

$$\mathcal{H}_k \colon \forall i \in [\![k, kn]\!],\, \mathbb{P}(S_k = i) = \frac{1}{n^k}\binom{i-1}{k-1}$$

À $n = 1$, on a bien, par équiprobabilité : $\forall i \in [\![1, n]\!], \mathbb{P}(S_1 = i) = \mathbb{P}(X_1 = i) = \frac{1}{n} = \frac{1}{n}\binom{i-1}{i-1}$. La propriété est initialisée. On suppose maintenant que $\mathcal{H}_k$ est vérifiée, pour $k \in [\![1, n-1]\!]$. On a alors, d'après la conclusion de la question 6b :

$$\forall i \in [\![k, n]\!], \mathbb{P}(S_{k+1} = i) = \frac{1}{n}\sum_{j=k}^{i-1} \mathbb{P}(S_k = j) = \frac{1}{n}\sum_{j=k}^{i-1} \frac{1}{n^k}\binom{j-1}{k-1} = \frac{1}{n^{k+1}}\sum_{j=k}^{i-1} \binom{j-1}{k-1} = \frac{1}{n^{k+1}}\binom{j}{k-1}$$

On a donc $\mathcal{H}_k \Rightarrow \mathcal{H}_{k+1}$. On conclut à partir du principe de raisonnement par récurrence :

$$\forall k \in [\![1, n]\!],\, \forall i \in [\![k, kn]\!],\, \mathbb{P}(S_k = i) = \frac{1}{n^k}\binom{i-1}{k-1}$$

8. (a) À partir de l'énoncé, on a $(T_n = k) = (S_k \geqslant n) \cap (S_{k-1} < n)$. De là, si $T_n > k$, alors $T_n \neq k$ et par négation $S_k < n$ car sinon, $S_{k-1} \geqslant n$, et alors $T_n \leqslant k - 1$, ce qui est absurde. Par ailleurs, si $S_k < n$, alors par définition $T_n > k$. On établit finalement :

$$(T_n > k) = (S_k < n) \Longrightarrow \mathbb{P}(T_n = k) = \mathbb{P}(S_k < n)$$

(b) Par l'égalité précédente :

$$\mathbb{P}(T_n > k) = \mathbb{P}(S_k < n) = \sum_{j=k}^{n-1} \mathbb{P}(S_k = j) = \sum_{j=k}^{n-1} \frac{1}{n^k}\binom{j-1}{k-1} = \frac{1}{n^k}\sum_{j=k}^{n-1}\binom{j-1}{k-1} = \frac{1}{n^k}\binom{n-1}{k}$$

9. On utilise la définition de l'espérance dans laquelle on fait apparaître une somme triangulaire :

$$\mathbb{E}(T_n) = \sum_{k=1}^{n} k\mathbb{P}(T_n = k) = \sum_{k=1}^{n}\sum_{j=1}^{k} \mathbb{P}(T_n = k) = \sum_{j=1}^{n}\underbrace{\sum_{k=j}^{n} \mathbb{P}(T_n = k)}_{\mathbb{P}(T_n \geqslant j)} = \sum_{j=1}^{n} \mathbb{P}(T_n > j - 1) = \sum_{k=0}^{n-1} \mathbb{P}(T_n > k)$$

On peut alors injecter l'expression obtenue en 8b et on obtient :

$$\mathbb{E}(T_n) = \sum_{k=0}^{n-1} \frac{1}{n^k}\binom{n-1}{k} = \sum_{k=0}^{n-1}\binom{n-1}{k}\left(\frac{1}{n}\right)^k \cdot 1^{n-1-k} = \left(\frac{1}{n} + 1\right)^{n-1}$$

en utilisant le binôme de Newton.

10. Étudions maintenant la limite de cette expression. On utilisera des équivalents.

$$(n-1)\ln\left(1 + \frac{1}{n}\right) \underset{\infty}{\sim} n \cdot \frac{1}{n} \longrightarrow 1$$

$$\Longrightarrow \left(1 + \frac{1}{n}\right)^{n-1} = \exp\left[(n-1)\ln\left(1 + \frac{1}{n}\right)\right] \underset{n\to\infty}{\longrightarrow} \exp 1$$

$$\therefore \lim_{n\to\infty} \mathbb{E}(T_n) = e$$