Des boules et des urnes

Préliminaires

Soit n un entier non nul et X une variable aléatoire suivant une loi uniforme sur $\{1, 2, \ldots, n\} = [\![1, n]\!]$.

1. Donner, sans justifications, l'espérance de X.
2. Prouver par récurrence que $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.
3. En déduire la variance de X

On considère une urne contenant N_1 boules blanches et N_2 boules noires indiscernables au toucher.

On pose $N = N_1 + N_2$.

On répète l'expérience suivante : on tire au hasard une boule dans l'urne et l'on replace dedans deux boules de la couleur obtenue.

À l'issue de la première expérience, l'urne contient donc $N + 1$ boules et l'on note X_1 la variable aléatoire égale eu nombre de boules blanches présentes dans l'urne. À l'issue de la deuxième expérience, l'urne contient donc $N + 2$ boules et l'on note X_2 la variable aléatoire égale au nombre de boules blanches dans l'urne.

Plus généralement, pour tout entier naturel k non nul, on note X_k la variable aléatoire égale au nombre de boules blanches présentes dans l'urne à l'issue de la k-ième expérience.

Pour tout k non nul, on note B_k l'événement "la boule tirée lors de la k-ième expérience est blanche".

Étude d'un cas particulier

On suppose ici que $N_1 = N_2 = 1$.

1. Déterminer la loi de X_1.
2. Déterminer la loi de X_2.
3. Soit $n \in \mathbb{N}^*$. Prouver que X_n suit une loi uniforme sur $[\![1, n]\!]$.
   On pourra faire une récurrence et utiliser le système complet $((X_n = k))_{1 \leqslant k \leqslant n+1}$ pour déterminer la loi de X_{n+1}.
4. Soit $n \in \mathbb{N}^*$. Déterminer la probabilité de B_{n+1}.
   On pourra utiliser la question précédente et la formule des probabilités totales.
5. Pour tout entier n non nul, on considère la variable aléatoire $Y_n = \frac{X_n - 1}{n}$.
   (a) Soit $n \in \mathbb{N}^*$. Donner la loi de Y_n.
   (b) On considère F la fonction de répartition d'une loi unifrme sur $[0, 1]$. Rappeler, pour tout réel x, la valeur de $F(x)$.
   (c) Soit $x \in [0, 1]$.
   Prouver que, pour tout entier n, on a $P(Y_n \leqslant x) = \frac{1}{n+1} \lfloor nx + 1 \rfloor$, où l'on note $\lfloor \cdot \rfloor$ la partie entière.
   (d) Pour tout entier $n \in \mathbb{N}^*$, on note F_{Y_n} la fonction de répartition de Y_n.
   Déduire de ce qui précède, que pour tout réel x, on a $\lim_{n \to +\infty} F_{Y_n}(x) = F(x)$.

Retour au cas général

1. Déterminer la probabilité des évènements B_1 et B_2.
2. Soit $n \in \mathbb{N} \backslash \{0, 1\}$.
   (a) Montrer que $\sum_{k=N_1}^{N_1+n-1} kP(X_{n-1} = k) = (N + 1n - 1)P(B_n)$.
   (b) Soit $k \in [\![N_1, N_1 + n - 1]\!]$.
   Déterminez la probabilité de B_{n+1} sachant $B_n \cap (X_{n-1} = k)$ puis la probabilité de B_{n+1} sachant $\overline{B_n} \cap (X_{n-1} = k)$.
   (c) En déduire que $P(B_{n+1}) = P(B_n)$.
3. Soit $n \in \mathbb{N}^*$. Déduire de la question précédente la probabilité de B_n et l'espéraance de X_n.

Corrigé

Préliminaires

1. L'espérance de X est $E(X) = \frac{n+1}{2}$.
2. Soit P_n le propriété à montrer. P_1 est évidente, et si P_n est vraie, on a

$$\begin{aligned}\sum_{k=1}^{n+1} k^2 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= (n+1)\left(\frac{n(2n+1)}{6} + n + 1\right) \\ &= (n+1)\left(\frac{2n^2+n+6n+6}{6}\right) \\ &= \frac{(n+1)(n+2)(2n+3)}{6}\end{aligned}$$

   et on a bien P_{n+1}. Par théorème de récurrence, on a bien l'égalité demandée.
3. On a donc, par lemme de transfert

$$E(X^2) = \sum_{k=0}^{n} k^2 P(X=k) = \frac{n(n+1)(2n+1)}{6n}.$$

   Par la formule de Koenig-Huygens, on a alors

$$V(X) = E(X^2) - E(x)^2 = \frac{n^2-1}{12}.$$

Étude d'un cas particulier

1. L'image de X_1 est $\{1,2\}$: si on a tiré une boule noire, alors $X_1 = 1$ (probabilité $1/2$), et si on a tiré une boule blanche, alors $X_1 = 2$ (probabilité $1/2$).
   X_1 suit donc une loi uniforme sur $[\![1,2]\!]$.
2. On note que l'image de X_2 est $\{1,2,3\}$. On a alors

$$\begin{aligned}P(X_2=1) &= P(\overline{B_1} \cap \overline{B_2}) \\ &= P(\overline{B_1}) P_{\overline{B_1}}(\overline{B_2}) \\ &= \frac{1}{2}\frac{2}{3} \\ &= \frac{1}{3}\end{aligned}$$

   De la même façon, on trouve $P(X_2 = 3) = P(B_1 \cap B_2) = \frac{1}{3}$, et on en déduit $P(X_2 = 2) = \frac{1}{3}$.
   X_2 suit donc une loi uniforme sur $[\![1,3]\!]$.
3. Soit P_n la proposition "$X_n \hookrightarrow \mathcal{U}(n+1)$". On a déjà montré P_1.
   Soit donc $n \in \mathbb{N}^*$, et supposons P_n. Comme le nombre de boules blanches reste le même ou augmente de 1, l'image de X_{n+1} est $[\![1,n+2]\!]$.
   Soit donc $k \in [\![1,n+2]\!]$. L'ensemble $\{[X_n = \ell] \mid \ell \in [\![1,n+1]\!]\}$ est un système complet d'événements, et donc on a

$$P(X_{n+1}=k) = \sum_{\ell=0}^{n+1} P(X_n=\ell) P_{[X_n=\ell]}(X_{n+1}=k) = \frac{1}{n+1}\sum_{\ell=1}^{n+1} P_{[X_n=\ell]}(X_{n+1}=k).$$

   On note que dans la somme, la plupart des termes sont nuls :

   - pour $k = 1$: pour avoir une boule blanche après le $n+1$-ième tirage, il faut qu'il y en ait une après le n-ième et qu'on ait tiré une boule noire (avec probabilité $\frac{n+1}{n+2}$). On a donc $P_{[X_n=\ell]}(X_{n+1}=k) = 0$ si $\ell > 1$, et donc

$$P(X_{n+1}=1) = \frac{1}{n+1}\frac{n+1}{n+2} = \frac{1}{n+2}.$$

- pour $k = n+2$: de la même façon, on a nécessairement $X_n = n+1$, et donc

$$P(X_{n+1} = n+2) = \frac{1}{n+2}.$$

- pour $k \in [\![2, n+1]\!]$: on a nécessairement $X_n = k$ ou $X_n = k-1$, et donc toutes les probabilités $P_{[X_n=\ell]}(X_{n+1} = k)$ pour $\ell \neq k, k-1$ sont nulles. On a donc

$$P(X_{n+1} = k) = \frac{1}{n+1}\left(P_{[X_n=k-1]}(X_{n+1} = k) + P_{[X_n=k]}(X_{n+1} = k)\right).$$

  Dans le premier cas, on a une urne avec $k-1$ boules blanches, et donc $P_{[X_n=k-1]}(X_{n+1} = k) = \frac{k-1}{n+2}$.
  Dans le second cas, on a une urne avec $n+2-k$ boules noires, et donc $P_{[X_n=k]}(X_{n+1} = k) = \frac{n+2-j}{n+2}$.
  Finalement, $P(X_{n+1} = k) = \frac{1}{n+2}$.

On a donc bien $X_{n+1} \hookrightarrow \mathcal{U}(n+2)$, et par théoreme de récurrence, on a donc bien

$$\forall n \in \mathbb{N}^*,\ X_n \hookrightarrow \mathcal{U}(n+1).$$

4. On a par la formule des probabilités totales

$$\begin{aligned}
P(B_{n+1}) &= \sum_{k=1}^{n+1} P(X_n = k) P_{[X_n=k]}(B_{n+1}) \\
&= \frac{1}{n+1}\sum_{k=1}^{n+1} P_{[X_n=k]}(B_{n+1}) \\
&= \frac{1}{n+1}\sum_{k=1}^{n+1} \frac{k}{n+2} \\
&= \frac{1}{(n+1)(n+2)} \frac{(n+1)(n+2)}{2} \\
&= \frac{1}{2}
\end{aligned}$$

5. (a) Il est clair que Y_n suit la loi uniforme sur $\left\{\frac{k}{n} \mid k \in [\![0, n]\!]\right\}$.

   (b) On a

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases}.$$

   (c) On a

$$\begin{aligned}
P(Y_n \leqslant x) &= P(X_n \leqslant nx+1) \\
&= \sum_{k=1}^{\lfloor nx+1 \rfloor} P(X_n = k) \\
&= \frac{\lfloor nx+1 \rfloor}{n+1}
\end{aligned}$$

   (d) Si $x < 0$ ou $x > 1$, il est clair que $F_{Y_n}(x) \to F(x)$.
   Soit donc $x \in [0,1]$. Prouvons que $\lfloor nx+1 \rfloor \sim nx$. On a

$$nx < \lfloor nx+1 \rfloor \leqslant nx+1,$$

   et donc

$$1 < \frac{\lfloor nx+1 \rfloor}{nx} \leqslant 1 + \frac{1}{nx}.$$

   Par encadrement des limites, on a bien l'équivalent cherché, et donc la limite voulue.

Retour au cas général

1. On a $P(B_1) = \frac{N_1}{N}$, et
$$P(B_2) = P(B_1)P_{B_1}(B_2) + P(\overline{B_1})P_{\overline{B_1}}(B_2) = \frac{N_1}{N}.$$

2. (a) On note que l'image de X_{n-1} est $[\![N_1, N_1+n-1]\!]$. Par la formule des probabilités totales, on a donc
$$\begin{aligned} P(B_n) &= \sum_{k=N_1}^{N_1+n-1} P(X_{n-1}=k)P_{[X_{n-1}=k]}(B_n) \\ &= \sum_{k=N_1}^{N_1+n-1} P(X_{n-1}=k)\frac{k}{N+n-1} \end{aligned}$$
On retrouve bien le résultat demandé.

   (b) On a
$$P_{B_n\cap[X_{n-1}=k]}(B_{n+1}) = \frac{k+1}{N+n}$$
et
$$P_{\overline{B_n}\cap[X_{n-1}=k]}(B_{n+1}) = \frac{k}{N+n}$$

   (c) On applique la formule des probabilités totales avec le système complet d'événements
$$\{[X_{n-1}=k]\cap B_n \mid k\in[\![N_1, N_1+n-1\}\cup\{[X_{n-1}=k]\cap\overline{B_n} \mid k\in[\![N_1, N_1+n-1\}.$$
$$\begin{aligned} P(B_n+1) &= \sum_{k=N_1}^{N_1+n-1} P([X_{n-1}=k]\cap B_n)P_{[X_{n-1}=k]\cap B_n}(B_{n+1}) + P([X_{n-1}=k]\cap\overline{B_n})P_{[X_{n-1}=k]\cap\overline{B_n}}(B_{n+1}) \\ &= \sum_{k=N_1}^{N_1+n-1} P([X_{n-1}=k]\cap B_n)\frac{k+1}{N+n} + P([X_{n-1}=k]\cap\overline{B_n})\frac{k}{N+n} \\ &= \frac{1}{N+n}\sum_{k=N_1}^{N_1+n-1} (k+1)P(X_{n-1}=k)P_{[X_{n-1}=k]}(B_n) + kP(X_{n-1}=k)P_{[X_{n-1}=k]}(\overline{B_n}) \\ &= \frac{1}{N+n}\sum_{k=N_1}^{N_1+n-1} P(X_{n-1}=k)\left(k(P_{[X_{n-1}=k]}(B_n) + P_{[X_{n-1}=k]}(\overline{B_n})) + P_{[X_{n-1}=k]}(B_n)\right) \\ &= \frac{1}{N+n}\sum_{k=N_1}^{N_1+n-1} P(X_{n-1}=k)\left(k+\frac{k}{N+n-1}\right) \\ &= \frac{1}{N+n}\left(1+\frac{1}{N+n-1}\right)\frac{1}{N+n}\sum_{k=N_1}^{N_1+n-1} kP(X_{n-1}=k) \\ &= P(B_n) \end{aligned}$$

3. On a donc bien pour tout n, $P(B_n) = \frac{N_1}{N}$. L'espérance de X_n est donnée par
$$E(X_n) = (N+n)P(B_{n+1}) = \frac{(N+n)N_1}{N}.$$