# RMS 2022 [984]

> $\boxed{\textbf{1}}$ *Une urne contient $n$ boules blanches et $n$ boules noires. On effectue des tirages sans remise (une boule à la fois) jusqu'à ce que les boules restant dans l'urne soient toutes de la même couleur.*
> *Modéliser la loi du nombre $X_n$ de boules qui restent dans l'urne à l'issue des tirages.*
> $\boxed{\textbf{2}}$ *On considère deux urnes contenant chacune $n$ boules. On effectue des tirages sans remise (une boule à la fois), en choisissant à chaque fois l'une des deux urnes de manière équiprobable jusqu'à ce que l'on constate que l'urne choisie soit vide.*
> *Modéliser la loi du nombre $Y_n$ de boules qui restent dans l'autre urne à ce moment-là. Calculer un équivalent de l'espérance de $Y_n$.*

$\boxed{\textbf{1}}$ On modélise les tirages par une famille $(U_k)_{1\leqslant k\leqslant 2n}$ de variables aléatoires de Bernoulli définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, en faisant l'hypothèse que

$$\forall\,(\varepsilon_1,\dots,\varepsilon_k)\in\{0,1\}^k,\quad \mathbf{P}(U_{k+1}=1 \mid U_1=\varepsilon_1,\dots,U_k=\varepsilon_k)=\frac{n-\sigma_k}{2n-k}$$

où $\sigma_k=(\varepsilon_1+\cdots+\varepsilon_k)$.

Ce modèle est raisonnable : si l'événement $[U_1=\varepsilon_1,\dots,U_k=\varepsilon_k]$ est réalisé, cela signifie qu'on a obtenu $\sigma_k=\varepsilon_1+\cdots+\varepsilon_k$ fois "1". On a donc obtenu $k-\sigma_k$ fois "0". Comme l'urne contient initialement $n$ boules de type "1" et $n$ boules de type "0", il reste donc $(2n-k)$ boules, parmi lesquelles on compte $(n-\sigma_k)$ boules de type "1" et (donc) $(n-k+\sigma_k)$ boules de type "0". Le quotient $(n-\sigma_k)/(2n-k)$ peut donc s'interpréter comme une hypothèse d'équiprobabilité sur le résultat du $(k+1)$-ième tirage.

En supposant de plus que

$$\mathbf{P}(U_1=1)=\frac{1}{2}\left(=\frac{n}{2n}\right),$$

ce modèle est complet, puisque les hypothèses qu'on vient de poser permettent de calculer

$$\mathbf{P}(U_1=\varepsilon_1,\dots,U_{2n}=\varepsilon_{2n})$$

quel que soit $(\varepsilon_1,\dots,\varepsilon_{2n})\in\{0,1\}^{2n}$ par l'intermédiaire de la Formule des probabilités composées.

❧ On peut en particulier vérifier que

$$\mathbf{P}(U_1=\varepsilon_1,\dots,U_{2n}=\varepsilon_{2n})=\frac{1}{\binom{2n}{n}}$$

pour toute famille $(\varepsilon_k)_{1\leqslant k\leqslant 2n}$ telle que $\sigma_{2n}=n$ (= les familles qui constituent le support de la loi du vecteur $(U_k)_{1\leqslant k\leqslant 2n}$).

Cette forme d'équiprobabilité va nous permettre de ramener le calcul de la loi de $X_n$ à du dénombrement.

► L'événement $[X_n=k]$ signifie qu'il ne reste $k$ boules, toutes du même type, dans l'urne. On a donc effectué $(2n-k)$ tirages, qui ont amené les $n$ boules d'un premier type et $(n-k)$ du second type :

$$[X_n=k]=[X_n=k,U_{2n}=1]\sqcup[X_n=k,U_{2n}=0]$$

et par équiprobabilité

$$\mathbf{P}(X_n=k)=2\,\mathbf{P}(X_n=k,U_{2n}=0).$$

► D'après la règle du jeu, l'événement $[X_n=k,U_{2n}=0]$ est égal à

$$[X_n=k,\underbrace{U_{2n-k}=1}_{\substack{\text{dernière boule}\\ \text{de type "1"}}},\underbrace{U_{(2n-k)+1}=U_{(2n-k)+2}=\cdots=U_{(2n-k)+k}=0}_{\text{les } k \text{ derniers tirages}}].$$

C'est donc la réunion de tous les événements

$$[U_1 = \varepsilon_1, \ldots, U_n = \varepsilon_n]$$

tels que $\varepsilon_{2n-k} = 1$ et $\varepsilon_{2n-k+1} = \varepsilon_{2n-k+2} = \cdots = \varepsilon_{2n} = 0$.

Comme il existe exactement $\binom{2n-k-1}{n-1}$ événements de ce genre (on place les $(n-1)$ boules de type "1" qui ont précédé la dernière boule de type "1" parmi les $(2n-k-1)$ premiers tirages), on en déduit que

$$\forall\, 1 \leqslant k \leqslant n, \quad \mathbf{P}(X_n = k) = \frac{2\binom{2n-k-1}{n-1}}{\binom{2n}{n}}.$$

❧ Il n'est pas inutile de vérifier la cohérence de ce résultat ! Comme

$$\begin{aligned}
\sum_{k=1}^{n} \binom{2n-k-1}{n-1} &= \sum_{k=1}^{n} \binom{(n-1)+(n-k)}{n-1} \\
&= \sum_{k=n-1}^{2n-2} \binom{k}{n-1} && (k \leftarrow 2n-k-1) \\
&= \binom{2n-1}{n} && \text{(triangle de Pascal)} \\
&= \frac{n}{2n}\binom{2n}{n} = \frac{1}{2} \cdot \binom{2n}{n}
\end{aligned}$$

on a bien $\mathbf{P}(X_n = 1) + \cdots + \mathbf{P}(X_n = n) = 1$.

**2❧** On modélise la seconde expérience par une famille $(V_k)_{1 \leqslant k \leqslant 2n+2}$ de variables aléatoires de Bernoulli définies sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ en supposant que la probabilité

$$\mathbf{P}(V_1 = \varepsilon_1, \cdots, V_{2n+2} = \varepsilon_{2n+2})$$

est la même pour tout vecteur $\boldsymbol{\varepsilon} = (\varepsilon_k)_{1 \leqslant k \leqslant 2n+2} \in \{0, 1\}^{2n+2}$ tel que

$$\sigma_{2n+2} = \sum_{k=1}^{2n+2} \varepsilon_k = n+1.$$

Cela revient à dire qu'on effectue de manière équiprobable $(2n+2)$ tirages successifs, tantôt dans l'urne "1", tantôt dans l'urne "0", et qu'on effectue exactement $(n+1)$ tirages dans chacune des urnes.

Le cardinal de l'ensemble $E_{n+1}$ des vecteurs $\varepsilon$ tels que $\sigma_{2n+2} = n+1$ est égal à $\binom{2n+2}{n+1}$ (on choisit les positions de $n$ éléments égaux à "0" parmi $(2n+2)$ positions possibles, les $n$ éléments égaux à "1" occuperont les positions restées libres).

► Ce modèle est évidemment complet, puisque la loi du vecteur aléatoire $(V_k)_{1 \leqslant k \leqslant 2n+2}$ est parfaitement définie.

► Ce modèle est aussi raisonnable au regard de la situation étudiée.

En effet, dans l'énoncé, chaque urne contient $n$ boules et après avoir tiré les $n$ boules de l'urne, il faut un $(n+1)$-ième tirage pour constater que l'urne est vide.

On peut aussi bien imaginer que chaque urne contient en fait $(n+1)$ boules et que la procédure s'arrête lors du tirage de cette $(n+1)$-ième boule fictive : cet artifice permet de ramener cette nouvelle situation à la situation précédente.

► Il reste cependant à vérifier que ce modèle est cohérent avec l'énoncé qui exige que

$$\forall\, 1 \leqslant k \leqslant 2n+2, \qquad \mathbf{P}(V_k = 1) = \frac{1}{2}.$$

Or

$$[V_k = 1] = \bigsqcup_{\substack{\varepsilon \in E_{n+1} \\ \text{t.q. } \varepsilon_k = 1}} [V_1 = \varepsilon_1, \ldots, V_{2n+2} = \varepsilon_{2n+2}].$$

Ces événements sont tous équiprobables et leur nombre est égal à $\binom{2n+1}{n}$ (il faut placer $n$ "1", la position $k$ est occupée par un "1" et on choisit $n$ positions parmi les $(2n+1)$ autres positions pour placer les $n$ autres "1"). Donc

$$\mathbf{P}(V_k = 1) = \frac{\binom{2n+1}{n}}{\binom{2n+2}{n+1}} = \frac{(2n+1)!}{n!(n+1)!} \cdot \frac{(n+1)!(n+1)!}{(2n+2)!} = \frac{n+1}{2n+2} = \frac{1}{2}.$$

Jusqu'ici, tout va bien...

► La valeur prise par la variable aléatoire $Y_n$ est comprise entre $0$ (lorsqu'on vide la seconde urne avant de constater que la première urne vidée est bien vide) et $n$ (lorsqu'on constate qu'une urne est vide avant de tirer la première boule de l'autre urne).

► Comme dans le premier cas,

$$[Y_n = k] = [Y_n = k, V_{2n+2} = 0] \sqcup [Y_n = k, V_{2n+2} = 1]$$

et par équiprobabilité,

$$\mathbf{P}(Y_n = k) = 2\,\mathbf{P}(Y_n = k, V_{2n+2} = 0).$$

L'événement $[V_{2n+2} = 0]$ signifie que le dernier tirage a lieu dans l'urne "0" et donc qu'on a déjà constaté que l'urne "1" a été vidée.

La contrainte $[Y_n = k]$ signifie alors qu'on a tiré $k$ fois de suite dans l'urne "0" avant d'effectuer le $(2n+2)$-ième et dernier tirage (également dans l'urne "0"). Autrement dit, on s'intéresse aux événements

$$[(V_1, \dots, V_{2n+2}) = \varepsilon]$$

tels que

$$V_{2n+1-k} = 1, \quad V_{(2n+1-k)+1} = \cdots = V_{(2n+1-k)+k} = 0, \quad V_{2n+2} = 0$$

(on constate que l'urne "1" est vide ; on tire encore $k$ boules dans l'urne "0" et on constate pour finir que l'urne "0" est vide elle aussi).

Le nombre de ces événements est égal à $\binom{2n-k}{n}$ (on choisit la place des $n$ boules "1" parmi les $(2n-k)$ positions libres) et par conséquent

$$\forall\, 0 \leqslant k \leqslant n, \qquad \mathbf{P}(Y_n = k) = \frac{2\binom{2n-k}{n}}{\binom{2n+2}{n+1}}.$$

REMARQUE.— On constate que les variables aléatoires $Y_n$ et $X_{n+1} - 1$ suivent la même loi : les encadrements $1 \leqslant k \leqslant n+1$ et $0 \leqslant k-1 \leqslant n$ sont équivalents et

$$\begin{aligned}\mathbf{P}(X_{n+1} - 1 = k-1) = \mathbf{P}(X_{n+1} = k) &= \frac{2\binom{2(n+1)-k-1}{(n+1)-1}}{\binom{2(n+1)}{n+1}} = \frac{2\binom{2n-(k-1)}{n}}{\binom{2n+2}{n+1}} \\ &= \mathbf{P}(Y_n = k-1).\end{aligned}$$

Si on veut traiter l'exercice dans le temps imparti, il serait judicieux de partir de cette remarque, en la fondant par un argument combinatoire (sans entrer dans les détails).

❧ Comme $Y_n$ est bornée, c'est bien une variable aléatoire d'espérance finie et, par définition de l'espérance,

$$\begin{aligned}\mathbf{E}(Y_n) &= \sum_{k=0}^{n} k\,\mathbf{P}(Y_n = k) \\ &= \frac{2}{\binom{2n+2}{n+1}} \sum_{k=0}^{n} k\binom{2n-k}{n} \\ &= \frac{2}{\binom{2n+2}{n+1}} \left[2n \sum_{k=n}^{2n} \binom{k}{n} - \sum_{k=n}^{2n} k\binom{k}{n}\right]. \qquad (k \leftarrow 2n-k)\end{aligned}$$

D'après la formule du triangle de Pascal (généralisée),

$$\sum_{k=n}^{2n} \binom{k}{n} = \binom{2n+1}{n+1} \quad \text{et} \quad \sum_{k=n}^{2n} k\binom{k}{n} = (n+1)\binom{2n+2}{n+2} - \binom{2n+1}{n+1}$$

donc

$$\forall\, n \geqslant 1, \qquad \mathbf{E}(Y_n) = \frac{2}{\binom{2n+2}{n+1}} \cdot \left[(2n+1)\binom{2n+1}{n+1} - (n+1)\binom{2n+2}{n+2}\right].$$

Après quelques simplifications,

$$\mathbf{E}(Y_n) = \frac{n}{n+2}$$

qui tend vers 1 lorsque $n$ tend vers $+\infty$.