Problème de Mathématiques

Référence pp1510 — Version du 31 décembre 2025

Soit n, un entier supérieur à 2.

Une urne contient $(n-1)$ boules blanches et une boule noire. On vide l'urne par des tirages successifs selon la règle suivante : les tirages d'ordre impair se font sans remise et les tirages d'ordre pair se font avec remise.

1. a. Quel est le nombre N de tirages effectués pour vider l'urne ?

1. b. Pour $1 \leqslant j < n$, combien reste-t-il de boules dans l'urne avant le $(2j)$-ième tirage ? avant le $(2j+1)$-ième tirage ?

Pour $1 \leqslant k \leqslant N$, on note X_k, la variable aléatoire qui décrit le résultat du k-ième tirage : elle vaut 1 si la boule noire apparaît et 0 si une boule blanche apparaît.

On désigne enfin par X, le nombre d'apparitions de la boule noire lors de l'ensemble des N tirages.

2. a. Calculer $\mathbf{P}(X_1 = 1)$ et $\mathbf{P}(X_2 = 1)$.

2. b. Pour $1 \leqslant j < n$, calculer $\mathbf{P}(X_{2j+1} = 1)$ et $\mathbf{P}(X_{2j} = 1)$.

2. c. En déduire la loi des variables X_k.

3. Pour $1 \leqslant j \leqslant n$, on note U_j, l'événement : *la boule noire apparaît pour la première fois lors du* $(2j-1)$-*ième tirage*.

3. a. En considérant l'état de l'urne avant le $(2n-2)$-ième tirage, démontrer que $\mathbf{P}(U_n) = 0$. Démontrer que

$$\forall\, 1 \leqslant j \leqslant n, \quad \mathbf{P}(U_j) = \frac{n-j}{n(n-1)}.$$

3. b. Exprimer l'événement $[X = 1]$ en fonction des événements U_j et en déduire la valeur de $\mathbf{P}(X = 1)$.

3. c. Démontrer que

$$\mathbf{P}(X = n) = \frac{1}{n!}.$$

4. En remarquant que

$$X = \sum_{k=1}^{2n-1} X_k,$$

calculer l'espérance de X.

5. Soit $0 \leqslant i \leqslant n-2$.

5. a. Calculer

$$\mathbf{P}(X_{2i+j+1} = 1 \mid X_{2i+1} = 1)$$

pour tout $1 \leqslant j \leqslant 2n-2i-2$.

5. b. En déduire que

$$\mathbf{Cov}(X_{2i+1}, X_{2i+j+1}) = \frac{-1}{n^2}.$$

6. Soit $1 \leqslant i < n$.

6. a. Démontrer que

$$\forall\, 1 \leqslant k < n-i, \quad \mathbf{P}(X_{2i+2k} = 1 \mid X_{2i} = 1) = \frac{1}{n-i}.$$

6. b. Démontrer que

$$\forall\, 0 \leqslant k < n-i, \quad \mathbf{P}(X_{2i+2k+1} = 1 \mid X_{2i} = 1) = \frac{1}{n-i}.$$

6. c. En déduire que

$$\forall\, 1 \leqslant j < 2n-2i, \quad \mathbf{Cov}(X_{2i}, X_{2i+j}) = \frac{i}{n^2(n-i)}.$$

7. Démontrer que

$$\mathbf{V}(X) = \frac{(2n+1)(n-1)}{n^2} - \frac{2}{n}\sum_{j=1}^{n-1}\frac{1}{j}.$$

Solution ❀ Tirages dans une urne

Il faut remarquer que le modèle probabiliste est seulement *suggéré* par l'énoncé. On doit donc admettre qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ sur lequel sont définies des variables aléatoires $X_1, \dots, X_N$ qui suivent des lois de Bernoulli et décrivent les résultats des tirages successifs.

Le modèle probabiliste est caractérisé par la loi de la famille $(X_1, \dots, X_N)$ et donc par la probabilité de chacun des événements
$$[X_1 = \varepsilon_1] \cap [X_2 = \varepsilon_2] \cap \cdots \cap [X_N = \varepsilon_N]$$
où $(\varepsilon_1, \dots, \varepsilon_N)$ parcourt $\{0, 1\}^N$.

On détermine les probabilités
$$\mathbf{P}(X_1 = \varepsilon_1, \dots, X_N = \varepsilon_N)$$
en se ramenant à des calculs de probabilités conditionnelles par la formule des probabilités composées :
$$\mathbf{P}(X_1 = \varepsilon_1) \prod_{i=2}^{N} \mathbf{P}(X_i = \varepsilon_i \mid X_1 = \varepsilon_1, \dots, X_{i-1} = \varepsilon_{i-1}).$$
On peut attribuer des valeurs raisonnables à $\mathbf{P}(X_1 = \varepsilon_1)$ et à
$$\mathbf{P}(X_i = \varepsilon_i \mid X_1 = \varepsilon_1, \dots, X_{i-1} = \varepsilon_{i-1})$$
en suivant un principe simple : comme on ne voit pas pourquoi chaque n'aurait pas la même probabilité d'être tirée, on supposera qu'à chaque tirage, chaque boule aura la même probabilité d'être tirée.

Conditionner par l'événement
$$[X_1 = \varepsilon_1, \dots, X_{i-1} = \varepsilon_{i-1}]$$
permet de connaître la composition de l'urne après les $(i-1)$ premiers tirages et donc d'en déduire, au moyen de notre hypothèse d'équiprobabilité, la probabilité conditionnelle de tirer une boule noire lors du i-ième tirage.

1. a. Pour vider l'urne, il faut retirer chacune des n boules qui y sont placées au début de l'expérience.

Les tirages de rang impair $(2k+1)$ (avec $k \geqslant 0$) enlèvent chacun *une* boule de l'urne ; les tirages de rang pair $2k$ (avec $k \geqslant 1$) ne modifient pas la composition de l'urne.

Il faut donc effectuer n tirages de rang impair, soit $0 \leqslant k < n$: la dernière boule est retirée lors du tirage de rang
$$2(n-1) + 1 = 2n - 1,$$
qui est donc le dernier tirage. Ainsi, $N = 2n - 1$.

1. b. Avant le $(2j)$-ième tirage, on a effectué les tirages de rang
$$1 = 2 \times 0 + 1, 2, 3 = 2 \times 1 + 1, \dots, 2j - 1 = 2(j-1) + 1.$$
On a donc effectué les tirages de rang impair $(2k+1)$ avec $0 \leqslant k < j$: autrement dit, on a effectué j tirages de rang impair, donc on a retiré exactement j boules.

Il reste $(n - j)$ boules avant le $(2j)$-ième tirage.

Comme le $(2j)$-ième tirage est un tirage de rang pair, il s'agit d'un tirage avec remise. Il reste donc encore $(n - j)$ boules avant le $(2j+1)$-ième tirage.

2. Comme les X_k sont, par hypothèse, des variables aléatoires, alors $[X_k = 0]$ et $[X_k = 1]$ sont des événements et pour alléger les notations, nous noterons respectivement
$$B_k \quad \text{et} \quad N_k \qquad \text{au lieu de} \qquad [X_k = 0] \quad \text{et} \quad [X_k = 1]$$
pour signifier qu'on tire une boule blanche (resp. noire) lors du k-ième tirage.

2. a. L'événement $[X_1 = 1]$ signifie qu'on tire une boule noire dans une urne constituée de n boules dont une seule est noire. Nous ferons donc l'hypothèse que
$$\mathbf{P}(X_1 = 1) = \frac{1}{n}.$$
❧ L'événement $[X_2 = 1]$ signifie qu'on retire une boule de l'urne avant de tirer une boule noire parmi les $(n-1)$ boules restantes : la boule retirée doit être blanche. En termes ensemblistes,
$$[X_2 = 1] \subset [X_1 = 0]$$

et par conséquent
$$[X_2 = 1] = [X_1 = 0] \cap [X_2 = 1].$$
D'après la formule des probabilités totales et l'hypothèse précédente sur la loi de X_1,
$$\begin{aligned}\mathbf{P}(X_2 = 1) &= \mathbf{P}(X_1 = 0)\,\mathbf{P}(X_2 = 1 \mid X_1 = 0)\\ &= \Big(1 - \frac{1}{n}\Big)\,\mathbf{P}(X_2 = 1 \mid X_1 = 0).\end{aligned}$$
Sachant qu'on a retiré une boule blanche, on tire une boule dans une urne contenant $(n-1)$ boules dont une seule est noire et nous ferons donc l'hypothèse (cohérente avec l'hypothèse précédente) que
$$\mathbf{P}(X_2 = 1 \mid X_1 = 0) = \frac{1}{n-1}.$$
Par conséquent,
$$\mathbf{P}(X_2 = 1) = \frac{n-1}{n} \cdot \frac{1}{n-1} = \frac{1}{n}.$$
2. b. Si on tire une boule noire lors du $(2j)$-ième ou du $(2j+1)$-ième tirage, alors il faut avoir tiré des boules blanches lors des tirages de rang $(2k+1)$ pour $0 \leqslant k < j$: autrement dit, les événements N_{2j} et N_{2j+1} sont contenus dans l'événement
$$A_j = \bigcap_{k=0}^{j-1} B_{2k+1}$$
donc $N_{2j} = A_j \cap N_{2j}$ et $N_{2j+1} = A_j \cap N_{2j+1}$.

❧ Si l'événement A_j est réalisé, on a retiré j boules blanches de l'urne, qui contient donc $(n-j)$ boules dont une seule est noire. La composition de l'urne reste la même lors du $(2j)$-ième tirage et du $(2j+1)$-ième tirage, donc
$$\mathbf{P}(N_{2j} \mid A_j) = \mathbf{P}(N_{2j+1} \mid A_j) = \frac{1}{n-j}$$
toujours pour rester cohérent avec notre hypothèse d'équiprobabilité des tirages.

❧ Il nous reste à calculer $\mathbf{P}(A_j)$ à l'aide de la formule des probabilités composées.

On a dit pourquoi $\mathbf{P}(B_0) = \mathbf{P}(X_1 = 0) = (n-1)/n$. Pour $1 \leqslant k < j$,
$$\mathbf{P}(B_{2k+1} \mid B_0 \cap \cdots B_{2k-1}) = \mathbf{P}(B_{2k+1} \mid A_k) = \frac{n-k-1}{n-k}$$
puisque l'urne contient $(n-k)$ boules avant le $(2k+1)$-ième tirage et que, sachant qu'on a retiré k boules blanches, il y reste une boule noire et $(n-k-1)$ boules blanches.

On en déduit que
$$\mathbf{P}(A_j) = \frac{n-1}{n}\prod_{k=1}^{j-1}\frac{n-(k+1)}{n-k} = \frac{n-j}{n}$$
et finalement que
$$\mathbf{P}(X_{2j+1} = 1) = \mathbf{P}(X_{2j} = 1) = \frac{1}{n-j} \cdot \frac{n-j}{n} = \frac{1}{n}.$$
2. c. D'après ce qui précède, les variables X_k suivent toutes la loi de Bernoulli de paramètre $1/n$.

3. a. Avant le $(2n-2)$-ième tirage, il ne reste plus qu'une seule boule $(n-(n-1)=1)$. De deux choses l'une :

— Si cette boule est blanche, c'est qu'on a déjà tiré la boule noire (à un tirage de rang impair) ;

— Si cette boule est noire, elle sera tirée lors du $(2n-2)$-ième tirage.

Dans les deux cas, la boule noire est tirée de l'urne *avant* le $(2n-1)$-ième (et dernier) tirage, donc $U_n = \varnothing$ et par conséquent $\mathbf{P}(U_n) = 0$.

❧ Soit maintenant $1 \leqslant j \leqslant n$. Par définition de U_j,
$$U_j = B_1 \cap B_2 \cap \cdots \cap B_{2j-2} \cap N_{2j-1}$$
(on obtient une boule blanche lors des $(2j-2)$ premiers tirages et une boule noire *pour la première fois* lors du $(2j-1)$-ième tirage) et nous allons encore appliquer la formule des probabilités composées.

On sait que
$$\mathbf{P}(B_1) = \frac{n-1}{n}.$$

Par un raisonnement analogue à celui du **2.b.**,

$$\mathbf{P}(N_{2j-1} \mid B_1 \cap B_2 \cap \cdots \cap B_{2j-3} \cap B_{2j-2}) = \frac{1}{n-(j-1)}$$

(il s'agit de tirer une boule noire dans une urne qui ne contient plus que $[n-(j-1)]$ boules dont une seule est noire) et, pour $1 \leqslant k < j$,

$$\mathbf{P}(B_{2k} \mid B_1 \cap B_2 \cap \cdots \cap B_{2k-1}) = \frac{n-k-1}{n-k}$$
$$\mathbf{P}(B_{2k+1} \mid B_1 \cap B_2 \cap \cdots \cap B_{2k-1} \cap B_{2k-2}) = \frac{n-k-1}{n-k}$$

(il s'agit dans les deux cas de tirer une boule blanche dans une urne qui ne contient plus que $(n-k)$ boules dont une seule n'est pas blanche). On en déduit que $\mathbf{P}(U_j)$ est égale à

$$\mathbf{P}(B_1)\Big[\prod_{k=1}^{j-2}\Big(\frac{n-k-1}{n-k}\Big)^2\Big]\frac{n-(j-1)-1}{n-(j-1)}\frac{1}{n-(j-1)}$$

donc (simplification télescopique)

$$\mathbf{P}(U_j) = \frac{1}{n}\cdot\frac{1}{n-1}\cdot(n-j+1-1) = \frac{n-j}{n(n-1)}.$$

3. b. *On n'a pas démontré que* X *était bien une variable aléatoire, donc on ne sait pas encore que* $[X=1]$ *est un événement. On va démontrer cette propriété en exprimant* $[X=1]$ *à l'aide des événements* U_j.

Comme on vide l'urne, la boule noire est tirée au moins une fois (lorsqu'on la retire de l'urne). La fonction X prend la valeur 1 lorsque la boule noire est tirée une seule fois, ce qui signifie qu'elle est tirée pour la première (et dernière) fois lors d'un tirage de rang *impair* (sitôt tirée, sitôt retirée de l'urne).

Autrement dit,

$$[X=1] = U_1 \sqcup U_2 \sqcup \cdots \sqcup U_{n-1},$$

ce qui prouve au passage que $[X=1] \in \mathcal{A}$, et par σ-additivité de $\mathbf{P}$,

$$\mathbf{P}(X=1) = \sum_{j=1}^{n-1}\mathbf{P}(U_j) = \sum_{j=1}^{n-1}\frac{n-j}{n(n-1)} = \sum_{k=1}^{n-1}\frac{k}{n(n-1)}$$

donc

$$\mathbf{P}(X=1) = \frac{1}{2}.$$

3. c. La fonction X prend la valeur n lorsque la boule noire est tirée n fois lors des $N = 2n+1$ tirages. Comme les n tirages de rang impair se font *sans* remise, cela signifie que la boule noire apparaît lors des $(n-1)$ tirages de rang pair et du dernier tirage de rang impair et que les boules blanches apparaissent lors des $(n-1)$ premiers tirages de rang impair. Autrement dit

$$[X=n] = B_1 \cap N_2 \cap B_3 \cap N_4 \cap \cdots \cap B_{2n-3} \cap N_{2n-2} \cap N_{2n-1}$$

ce qui prouve que $[X=n] \in \mathcal{A}$.

On reprend donc le raisonnement des **2.b.** et **3.a.** :

$$\mathbf{P}(B_1) = \frac{n-1}{n}$$
$$\mathbf{P}(N_{2j} \mid B_1 \cap N_2 \cap \cdots \cap N_{2j-2} \cap B_{2j-1}) = \frac{1}{n-j}$$
$$\mathbf{P}(B_{2j+1} \mid B_1 \cap N_2 \cap \cdots \cap B_{2j-1} \cap N_{2j}) = \frac{n-j-1}{n-j}$$
$$\mathbf{P}(N_{2n-1} \mid B_1 \cap N_2 \cap \cdots \cap B_{2n-3} \cap N_{2n-2}) = 1$$

d'où l'on déduit que $\mathbf{P}(X=n)$ est égale à

$$\frac{n-1}{n}\Big[\prod_{j=1}^{n-2}\frac{n-j-1}{(n-j)^2}\Big]\frac{1}{n-(n-1)}\cdot 1 = \frac{(n-1)!}{n[(n-1)!]^2}$$

c'est-à-dire
$$\mathbf{P}(X = n) = \frac{1}{n!}.$$

4. Les variables X_k sont des variables de Bernoulli (égales à 0 ou à 1), donc la somme des X_k est égale au nombre des X_k qui prennent la valeur 1, c'est-à-dire au nombre de fois où la boule noire a été tirée. Par conséquent,
$$X = \sum_{k=1}^{2n-1} X_k,$$
ce qui prouve (enfin !) que X est bien une variable aléatoire (en tant que somme de variables aléatoires).

Par linéarité de l'espérance,
$$\mathbf{E}(X) = \sum_{k=1}^{2n-1} \mathbf{E}(X_k) = \frac{2n-1}{n}$$
car chaque X_k suit la loi de Bernoulli de paramètre $1/n$.

5. a. Par **2.b.**, la probabilité de $[X_{2i+1} = 1]$ n'est pas nulle, donc on peut légitimement conditionner par cet événement.

Si cet événement est réalisé, la boule noire est retirée de l'urne après le $(2i+1)$-ième tirage et ne pourra donc plus être tirée par la suite. Donc
$$\forall\, 1 \leqslant j \leqslant 2n-2i-2, \quad [X_{2i+j+1} = 1] \cap [X_{2i+1} = 1] = \varnothing$$
(il faut que $2i + j + 1 \leqslant 2n - 1$) et la probabilité conditionnelle cherchée est donc nulle.

5. b. D'après la formule de Koenig-Huyghens,
$$\mathbf{Cov}(X_k, X_\ell) = \mathbf{E}(X_k X_\ell) - \mathbf{E}(X_k)\,\mathbf{E}(X_\ell).$$
D'après la question précédente, les deux variables de Bernoulli X_{2i+1} et X_{2i+j+1} ne peuvent être égales à 1 simultanément, donc le produit $X_{2i+1}X_{2i+j+1}$ est identiquement nul et
$$\mathbf{Cov}(X_{2i+1}, X_{2i+j+1}) = -\,\mathbf{E}(X_{2i+1})\,\mathbf{E}(X_{2i+j+1}) = \frac{-1}{n^2}$$
d'après **2.c.**

6. a. Sachant que l'événement $[X_{2i} = 1]$ est réalisé, on a effectué les $2i$ premiers tirages en ne retirant de l'urne que des boules blanches : on en a enlevé i, il reste donc $(n-i)$ boules dont une seule est noire.

Tout se passe donc comme si on calculait la probabilité d'obtenir une boule noire lors du $(2k)$-ième tirage dans une urne contenant initialement $(n-i-1)$ boules blanches et une boule noire. D'après **2.**,
$$\mathbf{P}(X_{2i+2k} = 1 \mid X_{2i} = 1) = \frac{1}{n-i}.$$

6. b. De même, tout se passe comme si on calculait la probabilité d'obtenir une boule noire lors du $(2k+1)$-ième tirage dans une urne contenant initialement $(n-i-1)$ boules blanches et une boule noire. On applique donc le résultat trouvé au **2.** en substituant $(n-i)$ à n.
$$\mathbf{P}(X_{2i+2k+1} = 1 \mid X_{2i} = 1) = \frac{1}{n-i}$$

6. c. Comme X_{2i} et X_{2i+j} sont des variables de Bernoulli, leur produit $X_{2i}X_{2i+j}$ est encore une variable de Bernoulli, donc
$$\mathbf{E}(X_{2i}X_{2i+j}) = \mathbf{P}(X_{2i}X_{2i+j} = 1) = \mathbf{P}(X_{2i} = 1, X_{2i+j} = 1)$$
et d'après la formule des probabilités totales et la question précédente,
$$\begin{aligned}\mathbf{E}(X_{2i}X_{2i+j}) &= \mathbf{P}(X_{2i} = 1)\,\mathbf{P}(X_{2i+j} = 1 \mid X_{2i} = 1)\\ &= \frac{1}{n}\cdot\frac{1}{n-i}.\end{aligned}$$
On déduit alors de la formule de Koenig-Huyghens que
$$\mathbf{Cov}(X_{2i}, X_{2i+j}) = \frac{1}{n(n-i)} - \frac{1}{n^2} = \frac{i}{n^2(n-i)}.$$

7. On sait calculer la variance d'une somme de variables aléatoires :

$$\mathbf{V}(X) = \sum_{k=1}^{2n-1} \mathbf{V}(X_k) + 2 \sum_{1 \leqslant j < k \leqslant 2n-1} \mathbf{Cov}(X_j, X_k).$$

Comme les X_k sont des variables de Bernoulli de paramètre $1/n$,

$$\begin{aligned}\sum_{k=1}^{2n-1} \mathbf{V}(X_k) &= (2n-1)\,\mathbf{V}(X_1) \\ &= (2n-1) \cdot \frac{1}{n} \cdot \Big(1 - \frac{1}{n}\Big) \\ &= \frac{(2n-1)(n-1)}{n^2}.\end{aligned}$$

On scinde la seconde partie de l'expression en sommant d'une part sur les indices j qui sont *pairs* ($j = 2i$ et $k = 2i + \ell$ pour $1 \leqslant i < n$) :

$$\sum_{1 \leqslant \ell < 2n-2i} \mathbf{Cov}(X_{2i}, X_{2i+\ell}) = \frac{(2n-2i-1)i}{n^2(n-i)}$$

et en sommant d'autre part sur les indices j qui sont *impairs* ($j = 2i+1$ et $k = 2i+\ell+1$ pour $0 \leqslant i \leqslant n-2$) :

$$\sum_{1 \leqslant \ell \leqslant 2n-2i-2} \mathbf{Cov}(X_{2i+1}, X_{2i+\ell+1}) = \frac{-(2n-2i-2)}{n^2}.$$

La contribution des indices j pairs est donc égale à

$$\begin{aligned}\sum_{i=1}^{n-1} \frac{(2n-2i-1)i}{n^2(n-i)} &= \frac{2}{n^2} \sum_{i=1}^{n-1} i - \frac{1}{n^2} \sum_{i=1}^{n-1} \frac{i}{n-i} \\ &= \frac{n-1}{n} - \frac{1}{n^2} \sum_{k=1}^{n} \frac{n-k}{k} \\ &= 1 - \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k}.\end{aligned}$$

La contribution des indices j impairs est égale à

$$\begin{aligned}\sum_{i=0}^{n-2} \frac{-(2n-2i-2)}{n^2} &= \frac{-2(n-1)^2}{n^2} + \frac{2}{n^2} \cdot \frac{(n-2)(n-1)}{2} \\ &= \frac{-(n-1)}{n} = -1 + \frac{1}{n}.\end{aligned}$$

La variance de X est donc égale à

$$\frac{(2n-1)(n-1)}{n^2} - \frac{2}{n} \sum_{k=2}^{n} \frac{1}{k}$$

(ne pas oublier le facteur 2 !), c'est-à-dire à

$$\frac{(2n+1)(n-1)}{n^2} - \frac{2}{n} \sum_{j=1}^{n-1} \frac{1}{j}$$

par une dernière application de l'astuce taupinale...