ESC 2005

Exercice 3

Sujet : Étude de la composition d'une urne évolutive — Moyen

Abordable en première année : ✗ — **Intérêt** : ★★★

Thèmes du programme abordés : variables aléatoires discrètes, espérance conditionnelle

Modifications apportées au sujet d'origine : ajout d'une question `Scilab`

Commentaires : un exercice bien guidé utilisant les espérances conditionnelles.

Dans tout l'exercice, S désigne un entier naturel non nul fixé.
Une urne contient initialement $4S$ boules indiscernables au toucher, dont S boules rouges, S boules vertes et $2S$ boules bleues.
On effectue des tirages successifs d'une boule, au hasard, et avec le protocole suivant :

- si la boule tirée est rouge, on ne la remet pas dans l'urne, mais on remet dans l'urne une boule bleue ;
- si la boule tirée est verte, on la remet dans l'urne ;
- si la boule tirée est bleue, on ne la remet pas dans l'urne, mais on remet dans l'urne une boule rouge.

On note, pour tout entier naturel n non nul, X_n la variable aléatoire égale au nombre de boules rouges présentes dans l'urne **après** le n-ème tirage, et on note X_0 la variable aléatoire certaine égale à S.
On rappelle que si A désigne un événement de probabilité non nulle et X une variable aléatoire discrète, $E(X|A)$ est l'espérance de X pour la probabilité conditionnelle P_A :

$$E(X|A) = \sum_{x \in X(\Omega)} x P_A(X = x).$$

1. Déterminer la loi de X_1, et calculer son espérance.
2. Déterminer la loi de X_2 et calculer son espérance.
3. On suppose désormais que n est un entier supérieur ou égal à $2S$, de sorte que

$$X_n(\Omega) = \{0, \ldots, 3S\}.$$

   a. Soit $k \in [\![1, 3S-1]\!]$.
   Quelle est la composition de l'urne une fois l'événement $[X_n = k]$ réalisé ?
   En déduire la loi de X_{n+1} conditionnellement à l'événement $[X_n = k]$.

   b. Montrer que $E(X_{n+1}|X_n = k) = \left(1 - \dfrac{1}{2S}\right)k + \dfrac{3}{4}$.
   Cette formule est-elle encore vraie pour $k = 0$ et $k = 3S$?

   c. En déduire par la formule de l'espérance totale que $E(X_{n+1}) = \left(1 - \dfrac{1}{2S}\right)E(X_n) + \dfrac{3}{4}$.

4. On note, pour tout entier naturel n supérieur ou égal à $2S$, $u_n = E(X_n)$.

   a. Déterminer le réel α tel que $\alpha = \left(1 - \dfrac{1}{2S}\right)\alpha + \dfrac{3}{4}$.

   b. Montrer que la suite $(u_n - \alpha)_{n \geqslant 2S}$ est géométrique.

   c. En déduire l'expression de u_n en fonction de n, S et u_{2S}.

   d. Montrer que $\lim\limits_{n \to \infty} E(X_n) = \dfrac{3S}{2}$.

5. Écrire en `Scilab` une fonction `simule(S,n)` qui prend en entrée deux paramètres S et n, simule une série de n tirages dans une urne de $4S$ boules et retourne le nombre de boules rouges restant dans l'urne à l'issue de ces n tirages.

ESC 2005 : CORRIGÉ

EXERCICE 3

Notons qu'à tout moment, puisque le nombre total de boules est $4S$ et que le nombre de boules vertes est S, on a toujours

$$\text{(nombre de boules rouges)} + \text{(nombre de boules bleues)} = 3S.$$

1. Puisque nous connaissons la composition initiale de l'urne, nous pouvons affirmer que $X_1(\Omega) = \{S-1, S, S+1\}$ et que

$$P(X_1 = S-1) = \frac{S}{4S} = \frac{1}{4}, P(X_1 = S) = \frac{1}{4} \text{ et } P(X_1 = S+1) = \frac{2S}{4S} = \frac{1}{2}.$$

On en déduit que

$$E(X_1) = \frac{S-1}{4} + \frac{S}{4} + \frac{S+1}{2} = \boxed{S + \frac{1}{4}}.$$

Détails
$X_1 = S-1$ si on a tiré une boule rouge et $X_1 = S+1$ si on a rajouté une boule rouge, c'est-à-dire tiré une boule bleue.

2. L'énoncé ne donnant aucune contrainte sur S (si ce n'est que $S \neq 0$), il va nous falloir traiter séparément les cas $S = 1$ et $S > 1$.
Si $S > 1$: alors $X_2(\Omega) = [\![S-2, S+2]\!]$.
En considérant le système complet d'événements $[X_1 = S-1], [X_1 = S], [X_1 = S+1]$, et en appliquant la formule des probabilités totales, il vient :

$$P(X_2 = S-2) = P_{[X_1=S-1]}(X_2 = S-2)P(X_1 = S-1)+P_{[X_1=S]}(X_2 = S-2)P(X_1 = S)+P_{[X_1=S+1]}(X_2 = S-2)P(X_1 = S+1).$$

Mais si l'urne contenait S ou $S+1$ boules rouges à l'issue du premier tirage, elle en comptera toujours au moins $S-1$ à l'issue du second, et jamais $S-2$, donc

$$P_{[X_1=S]}(X_2 = S-2) = P_{[X_1=S+1]}(X_2 = S+2) = 0.$$

Et donc

$$P(X_2 = S-2) = P_{[X_1=S-1]}(X_2 = S-2)P(X_1 = S-1) = \frac{S-1}{4S}\frac{1}{4} = \boxed{\frac{S-1}{16S}}.$$

De même, on a

$$\begin{aligned} P(X_2 = S-1) &= P_{[X_1=S-1]}(X_2 = S-1)P(X_1 = S-1) + P_{[X_1=S]}(X_2 = S-1)P(X_1 = S) \\ &= \frac{1}{4}\frac{1}{4} + \frac{1}{4}\frac{1}{4} = \boxed{\frac{1}{8}}. \end{aligned}$$

$$\begin{aligned} P(X_2 = S) &= P_{[X_1=S-1]}(X_2 = S)P(X_1 = S-1) + P_{[X_1=S]}(X_2 = S)P(X_1 = S) + P_{[X_1=S+1]}(X_2 = S)P(X_1 = S+1) \\ &= \frac{2S+1}{4S}\frac{1}{4} + \frac{1}{4}\frac{1}{4} + \frac{S+1}{4S}\frac{1}{2} = \boxed{\frac{5S+3}{16S}}. \end{aligned}$$

$$\begin{aligned} P(X_2 = S+1) &= P_{[X_1=S]}(X_2 = S+1)P(X_1 = S) + P_{[X_1=S+1]}(X_2 = S+1)P(X_1 = S+1) \\ &= \frac{1}{2}\frac{1}{4} + \frac{1}{4}\frac{1}{2} = \boxed{\frac{1}{4}}. \end{aligned}$$

Et enfin,

$$P(X_2 = S+2) = P_{[X_1=S+1]}(X_2 = S+2)P(X_1 = S+1) = \frac{2S-1}{4S}\frac{1}{2} = \boxed{\frac{2S-1}{8S}}.$$

Astuce
Apprenons à déceler à peu de frais les erreurs de calcul : la somme de ces probas doit valoir 1, prenez le temps de le vérifier au brouillon !

Si $S = 1$: alors X_2 ne peut prendre la valeur $S-2$ (car si $X_1 = S-1$, on ne peut plus tirer de boules rouges), mais seulement les valeurs $0, 1, 2, 3$. On a alors $P(X_2 = -1) = 0$ ce qui est cohérent avec la formule donnée précédemment pour $P(X_2 = S-1)$ dans le cas où $S = 1$.
Les formules donnant les autres probabilités sont inchangées par rapport au cas où $S > 1$.

Dans les deux cas, l'espérance de X_2 est donnée par

$$E(X_2) = (S-2)\frac{S-1}{16S} + (S-1)\frac{1}{8} + S\frac{5S+3}{16S} + \frac{S+1}{4} + (S+2)\frac{2S-1}{8S} = \frac{16S^2+8S-2}{16S} = \boxed{S + \frac{1}{2} - \frac{1}{8S}}.$$

3.a. Une fois l'événement $[X_n = k]$ réalisé, l'urne contient k boules rouges, S boules vertes[1], et $3S - k$ boules bleues en vertu de la remarque faite au début.
On a alors $P_{[X_n=k]}(X_{n+1} = k+1)$ qui est la probabilité d'avoir tiré une boule bleue lors du $(n+1)$-ième tirage, sachant que l'urne contenait k boules rouges (et donc $3S - k$ boules bleues). Ainsi

$$P_{[X_n=k]}(X_{n+1} = k+1) = \frac{3S-k}{4S}.$$

De même, $P_{[X_n=k]}(X_{n+1} = k) = \dfrac{S}{4S} = \dfrac{1}{4}$.
Et donc $P_{[X_n=k]}(X_{n+1} = k-1) = \dfrac{k}{4S}$.

[1] Le nombre de boules vertes ne change pas au cours du temps.

Remarque
On ne détermine que ces trois valeurs, car si $[X_n = k]$, alors X_{n+1} ne peut prendre que les valeurs $k-1$, k ou $k+1$ (on n'a ajouté ou retiré qu'au plus une boule).

3.b. Puisque nous avons déterminé la loi de X_{n+1} sachant $[X_n = k]$, nous pouvons calculer l'espérance conditionnelle de X_{n+1} sachant $[X_n = k]$:

$$E(X_{n+1}|X_n = k) = (k+1)\frac{3S-k}{4S} + \frac{k}{4} + (k-1)\frac{k}{4S} = k\frac{4S-2}{4S} + \frac{3S}{4S} = \boxed{\left(1 - \frac{1}{2S}\right)k + \frac{3}{4}.}$$

Pour $k = 0$, nous ne pouvons pas utiliser le résultat de la question précédente, mais nous savons qu'alors $P_{[X_n=0]}(X_{n+1} = 1) = \dfrac{3S}{4S} = \dfrac{3}{4}$, et que $P_{[X_n=0]}(X_{n+1} = 0) = \dfrac{1}{4}$. Donc

$$E(X_{n+1}|X_n = 0) = 1 \times \frac{3}{4} = \left(1 - \frac{1}{2S}\right) \times 0 + \frac{3}{4}.$$

Ainsi, $\boxed{\text{la formule reste valable pour } k = 0.}$

Pour $k = 3S$, là encore il faut refaire le calcul : on a $P_{[X_n=3S]}(X_{n+1} = 3S) = \dfrac{1}{4}$ et $P_{[X_n=3S]}(X_{n+1} = 3S-1) = \dfrac{3S}{4S} = \dfrac{3}{4}$. Alors

$$E(X_{n+1}|X_n = 3S) = (3S-1)\frac{3}{4} + \frac{3S}{4} = 3S - \frac{3}{4} = 3S - \frac{3}{2} + \frac{3}{4} = \left(1 - \frac{1}{2S}\right)3S + \frac{3}{4}.$$

Ainsi, $\boxed{\text{la formule est encore valable pour } k = 3S.}$

3.c. Puisque X_{n+1} est une variable finie, elle admet une espérance.
Par la formule de l'espérance totale appliquée au système complet d'événements $\{[X_n = k,\ k \in [\![0;3S]\!]\}$, on a alors

$$\begin{aligned}
E(X_{n+1}) &= \sum_{k=0}^{3S} E(X_{n+1}|X_n = k)P(X_n = k) \\
&= \sum_{k=0}^{3S} \left(\left(1 - \frac{1}{2S}\right)k + \frac{3}{4}\right)P(X_n = k) \\
&= \left(1 - \frac{1}{2S}\right)\sum_{k=0}^{3S} kP(X_n = k) + \frac{3}{4}\sum_{k=0}^{3S} P(X_n = k) \\
&= \left(1 - \frac{1}{2S}\right)E(X_n) + \frac{3}{4}\sum_{k=0}^{3S} P(X_n = k).
\end{aligned}$$

Mais, puisque $\{[X_n = k], k \in [\![0, 3S]\!]\}$ est un système complet d'événements, alors

$$\sum_{k=0}^{3S} P(X_n = k) = 1.$$

Et donc

$$\boxed{E(X_{n+1}) = \left(1 - \frac{1}{2S}\right)E(X_n) + \frac{3}{4}.}$$

Hypothèses
Si l'on a remarqué que X_{n+1} admet une espérance, alors l'énoncé de la formule de l'espérance totale nous garantit l'existence des espérances conditionnelles (déjà prouvée à la question 3.c) et la convergence de la somme (qui est en fait évidente car c'est une somme finie).
Si au contraire on ne voit pas que X_{n+1} admet automatiquement une espérance, il faut alors dire qu'elle existe car les espérances conditionnelles existent et que la somme est finie, donc convergente.

4.a. En résolvant l'équation, on obtient

$$\alpha = \left(1 - \frac{1}{2S}\right)\alpha + \frac{3}{4} \Leftrightarrow \alpha\frac{1}{2S} = \frac{3}{4} \Leftrightarrow \boxed{\alpha = \frac{3S}{2}.}$$

4.b. En utilisant la relation de la question 3.c, on a pour $n \geqslant 2S$,

$$u_{n+1}-\alpha = \left(1 - \frac{1}{2S}\right)u_n+\frac{3}{4}-\alpha = \left(1 - \frac{1}{2S}\right)(u_n-\alpha)+\underbrace{\left(1 - \frac{1}{2S}\right)\alpha + \frac{3}{4} - \alpha}_{=0} = \left(1 - \frac{1}{2S}\right)(u_n-\alpha).$$

Et donc $\boxed{\text{la suite } (u_n - \alpha)_{n\geqslant 2S} \text{ est géométrique de raison } \left(1 - \frac{1}{2S}\right).}$

Remarque

Vous avez probablement reconnu qu'en réalité, (u_n) est une suite arithmético-géométrique.

4.c. Par la question précédente, en posant $v_n = u_n - \frac{3S}{2}$, alors $(v_n)_{n\geqslant 2S}$ est une suite géométrique de raison $1 - \frac{1}{2S}$. Ainsi,

$$\forall n \geqslant 2S,\ v_n = v_{2S}\left(1 - \frac{1}{2S}\right)^{n-2S} v_{2S}.$$

Mais $v_{2S} = u_{2S} - \frac{3S}{2}$. Ainsi,

$$\boxed{\forall n \geqslant 2S,\ u_n = \left(1 - \frac{1}{2S}\right)^{n-2S}\left(u_{2S} - \frac{3S}{2}\right) + \frac{3S}{2}.}$$

Danger !

Ici, la suite n'est géométrique qu'à partir du rang $2S$, nous ne l'avons pas étudiée avant. La formule usuelle

$$v_n = v_0\left(1 - \frac{1}{2S}\right)^n$$

n'est donc pas valable ici !

4.d. Puisque $0 < 1 - \frac{1}{2S} < 1$, nous savons que $\lim\limits_{n\to\infty}\left(1 - \frac{1}{2S}\right)^{n-2S} = 0$ et on en déduit donc que

$$\boxed{\lim_{n\to\infty} E(X_n) = \lim_{n\to\infty} u_n = \frac{3S}{2}.}$$

Intuition

Les boules rouges et bleues jouent des rôles symétriques dans l'expérience, et il y en a $3S$ en tout.
Même si au début, les boules bleues sont en plus grand nombre, au bout d'un certain temps (c'est-à-dire pour n grand), il doit y avoir en moyenne autant de boules rouges que de boules bleues dans l'urne, c'est-à-dire $\frac{3S}{2}$.

5. Une option[2] est par exemple le programme suivant :

```
1  function y=simule(S,n)
2      y=S
3      A = grand(1,n,'uin',1,4*S)
4      for i=1 :n
5          t = A(i)
6          if t <= y then
7              y = y-1
8          else if t>y+S then
9                  y=y+1
10             end
11         end
12     end
13 endfunction
```

Expliquons rapidement ce programme. La variable y sert à compter le nombre de boules rouges. On assimile les n boules aux nombres de $[\![1, 4S]\!]$, en considérant que les y premiers correspondent au boules rouges, que les S suivants correspondent aux boules vertes, et que les $3S - y$ derniers sont les boules bleues.
Nous commençons par déterminer directement les numéros des n boules tirées à l'aide de la ligne 3.
Puis, à l'aide d'une boucle `for` on modifie la composition de l'urne pour chaque tirage :

- si la boule tirée est rouge (`if t<=y`), on enlève une boule rouge de l'urne (`y = y-1`)
- si la boule tirée est bleue, on ajoute une boule rouge (`y = y+1`).
- si la boule tirée est verte, c'est-à-dire si son numéro est dans $[\![y+1, y+S]\!]$, il ne se passe rien.

A l'issue de la boucle, y contient le nombre de boules rouges restant dans l'urne à l'issue du n-ième tirage.

[2] Parmi beaucoup d'autres !