Correction du DM 007

Exercice 1 : Lancers de boules dans des cases

1) Puisque les lancers sont successifs et indépendants, on peut considérer $\boxed{\Omega = [\![1, N]\!]^n}$ muni de $\mathcal{P}(\Omega)$ et de l'équiprobabilité. Si $\omega \in \Omega$ alors, pour tout $j \in [\![1, n[\![$, la $j^{\text{ième}}$ coordonnée de ω représente le numéro de la case dans laquelle se trouve la $j^{\text{ième}}$ boule. On a $\boxed{\mathrm{card}(\Omega) = N^n.}$

2) Au minimum, il y a une seule case atteinte (toutes les boules dans la même case) et, puisque $n \geqslant N$, au maximum, les N cases sont occupé.

   Pour tout $k \in [\![1, N[\![$, $\mathbb{P}(T_n = k) > 0$. En effet cette probabilité est minorée, par exemple, par la probabilité que les k premières boules aient atterris dans les k premières cases respectivement puis que les $n-k$ suivantes dans la première case (cette probabilité étant égale, par indépendance des lancers, à $\dfrac{1}{N^n} > 0$). Ainsi $\boxed{T_n(\Omega) = [\![1, N]\!].}$

3) *Bon d'accord l'énoncé dit que $2 \leqslant N \leqslant n$ donc T_1 n'existe pas et T_2 n'est pas très intéressante à étudier. Dans cette question, faisons comme si il n'y avait pas cette condition.*

   - T_1 est le nombre de cases atteintes après un seul lancer si bien que $\boxed{T_1 = \{1\} \text{ et } \mathbb{P}(T_1 = 1) = 1.}$
   - T_2 est le nombre de cases atteintes après deux lancers donc $\boxed{T_2 = \{1, 2\}.}$ Nous avons $\mathbb{P}(T_2 = 1) = \dfrac{\mathrm{card}([T_1 = 1])}{\mathrm{card}(\Omega)}$. Or $[T_2 = 1]$ est l'événement $\{(1,1), (2,2), \dots, (N,N)\}$ dont le cardinal est égale à N. Ainsi

   $$\boxed{\mathbb{P}(T_2 = 1) = \frac{N}{N^2} = \frac{1}{N} \qquad \text{et} \qquad \mathbb{P}(T_2 = 2) = 1 - \mathbb{P}(T_2 = 1) = \frac{N-1}{N}.}$$

4)
   - Réaliser l'événement $[T_n = 1]$ revient à choisir une case (N choix) puis à placer toutes les boules dans cette case (1 choix). Ainsi $\mathrm{card}([T_n = 1]) = N$ et donc $\boxed{\mathbb{P}(T_n = 1) = \dfrac{N}{N^n} = \dfrac{1}{N^{n-1}}.}$
   - Réaliser l'événement $[T_n = 2]$ revient à choisir deux cases ($\binom{N}{2}$ choix) puis à placer les boules au hasard dans ces deux cases ($2^n - 2$ choix : 2 choix pour chacune des n boules mais il faut exclure les deux configurations où les boules sont toutes dans une seule des deux cases). Ainsi $\mathrm{card}([T_n = 2]) = \binom{N}{2}(2^n - 2)$ et donc $\boxed{\mathbb{P}(T_n = 2) = \dfrac{N(N-1)(2^n-2)}{2\,N^n} = \dfrac{(N-1)(2^{n-1}-1)}{N^{n-1}}.}$

5) Soit $k \in [\![1, N]\!]$. Soit $n > N$. D'après la formule des probabilités totales appliquée au système complet d'événements associé à la variable aléatoire réelle finie T_n, on a

   $$\mathbb{P}(T_{n+1} = k) = \sum_{j=1}^{N} \mathbb{P}_{[T_n = j]}(T_{n+1} = k)\mathbb{P}(T_n = j).$$

   Soit $j \in [\![1, N]\!]$. Sachant que $T_n = j$, si on lance la $(n+1)^{\text{ième}}$ boule, alors

   - ou bien elle tombe dans une des j cases déjà occupées (ce qui arrive avec probabilité $\dfrac{j}{N}$) et $[T_{n+1} = j]$ est réalisé.
   - ou bien elle tombe dans une case jusqu'alors inoccupée (ce qui arrive avec probabilité $\dfrac{N-j}{N}$) et $[T_{n+1} = j+1]$ est réalisé.

Par conséquent $\mathbb{P}_{[T_n=k]}(T_{n+1}=k)=\dfrac{k}{N}$, $\mathbb{P}_{[T_n=k-1]}(T_{n+1}=k)=\dfrac{N-k+1}{N}$ et $\mathbb{P}_{[T_n=j]}(T_{n+1}=k)=0$ pour tout $j\in[\![1,N]\!]\backslash\{k,k-1\}$. Ainsi

$$\boxed{\mathbb{P}(T_{n+1}=k)=\frac{k}{N}\mathbb{P}(T_n=k)+\frac{N-k+1}{N}\mathbb{P}(T_n=k-1).}$$

6) a) Puisque $([T_n=k])_{1\leqslant k\leqslant N}$ est un s.c.e, on a $1=\sum_{k=1}^{N}\mathbb{P}(T_n=k)=\sum_{k=1}^{N}\mathbb{P}(T_n=k)1^k$, c'est-à-dire $\boxed{G_n(1)=1.}$

b) La fonction G_n est polynomiale donc elle est dérivable sur $\mathbb{R}$ et

$$\forall x\in\mathbb{R},\qquad G_n'(x)=\sum_{k=1}^{N}k\mathbb{P}(T_n=k)x^{k-1}.$$

Or Ainsi $\mathbb{E}(T_n)=\sum_{k=1}^{N}k\mathbb{P}(T_n=k)=\sum_{k=1}^{N}k\mathbb{P}(T_n=k)1^{k-1}$, c'est-à-dire $\boxed{G_n'(1)=\mathbb{E}(T_n).}$

c) Soit $x\in\mathbb{R}$. On a

$$G_{n+1}(x)=\sum_{k=1}^{N}\mathbb{P}(T_{n+1}=k)x^k=\sum_{k=1}^{N}\frac{k}{N}\mathbb{P}(T_n=k)x^k+\sum_{k=1}^{N}\frac{N-k+1}{N}\mathbb{P}(T_n=k-1)x^k.$$

Regardons chaque somme séparément.

- On a $\sum_{k=1}^{N}\dfrac{k}{N}\mathbb{P}(T_n=k)x^k=\dfrac{x}{N}\sum_{k=1}^{N}k\mathbb{P}(T_n=k)x^{k-1}=\dfrac{x}{N}G_n'(x)$.
- Faisons le changement de variable $j=k-1$ dans la deuxième somme :

$$\sum_{k=1}^{N}\frac{N-k+1}{N}\mathbb{P}(T_n=k-1)x^j=\sum_{j=0}^{N-1}\frac{N-j}{N}\mathbb{P}(T_n=j)x^{j+1}=\sum_{j=1}^{N}\frac{N-j}{N}\mathbb{P}(T_n=j)x^{j+1},$$

puisque $\mathbb{P}(T_n=0)=0$ et $\dfrac{N-N}{N}\mathbb{P}(T_n=N)=0$. Ensuite

$$\begin{aligned}\sum_{k=1}^{N}\frac{N-k+1}{N}\mathbb{P}(T_n=k-1)x^j&=\sum_{j=1}^{N}\mathbb{P}(T_n=j)x^{j+1}-\frac{1}{N}\sum_{j=1}^{N}j\mathbb{P}(T_n=j)x^{j+1}\\&=x\sum_{j=1}^{N}\mathbb{P}(T_n=j)x^j-\frac{x^2}{N}\sum_{j=1}^{N}j\mathbb{P}(T_n=j)x^{j-1}\\&=xG_n(x)-\frac{x^2}{N}G_n'(x).\end{aligned}$$

Nous en déduisons que

$$G_{n+1}(x)=\sum_{k=1}^{N}\mathbb{P}(T_{n+1}=k)x^k=\frac{x}{N}G_n'(x)+xG_n(x)-\frac{x^2}{N}G_n'(x).$$

et donc $\boxed{G_{n+1}(x)=\dfrac{1}{N}(x-x^2)G_n'(x)+xG_n(x).}$

d) Nous avons $\mathbb{E}(T_{n+1})=G_{n+1}'(1)$. Or, pour tout $x\in\mathbb{R}$,

$$G_{n+1}'(x)=\frac{(1-2x)}{N}G_n'(x)+\frac{x-x^2}{N}G_n''(x)+G_n(x)+xG_n'(x).$$

Ainsi $\mathbb{E}(T_{n+1})=\dfrac{-1}{N}G_n'(1)+\dfrac{0}{N}G_n''(1)+G_n(1)+G_n'(1)=\boxed{1+\left(1-\dfrac{1}{N}\right)\mathbb{E}(T_n)}.$

e) La suite $(\mathbb{E}(T_n))_{n\in\mathbb{N}}$ est arithmético-géométrique. Pour tout $x \in \mathbb{R}$, on a $x = 1 + \left(1 - \dfrac{1}{N}\right)x$ si et seulement si $x = N$. Nous en déduisons que la suite $(\mathbb{E}(T_n) - N)_{n\in\mathbb{N}}$ est géométrique de raison $1 - \dfrac{1}{N}$. Finalement

$$\forall n \in \mathbb{N}, \qquad \mathbb{E}(T_n) = N\left(1 - \left(1 - \frac{1}{N}\right)^n\right).$$

Puisque $\left|1 - \dfrac{1}{N}\right| < 1$, nous en déduisons que $\boxed{\lim\limits_{n\to+\infty} E(T_n) = N.}$

Exercice 2 : Loi des tirages avec ou sans remise

1) Les N tirages successifs sont des expériences de Bernoulli puisqu'il y a deux issues possibles : tirer une boule rouge (succès) ou une boule bleue (échec). Ils sont identiques et indépendants (puisqu'on effctue ces tirages avec remise). Le nombre R de boules rouges obtenues est donc le nombre de succès. Par conséquent $\boxed{R \hookrightarrow \mathcal{B}(N, p).}$

2) a) Puisque les tirages se font successivement et sans remise, on peut considérer Ω l'ensemble des tirages successifs et sans remise de n boules dans l'urne. On peut aussi supposer que les boules rouges sont numérotées de 1 à r et les boules bleues sont numérotées de $r+1$ à $r+b = N$ et considérer Ω l'ensemble des n-listes d'éléments de $[\![1, N]\!]$. On a alors

$$\boxed{\mathrm{card}(\Omega) = \frac{N!}{(N-n)!}.}$$

On munit $(\Omega, \mathcal{P}(\Omega))$ de l'équiprobabilité.

Remarque : *Les tirages se font successivement et sans remise mais, comme on ne s'intéresse qu'au nombre de boules rouges tirées, l'ordre des tirages ne compte pas. On peut donc supposer que l'on tire simultanément n boules dans l'urne considérer ainsi Ω l'ensemble des tirages simultanés. Dans ce cas $\mathrm{card}(\Omega) = \binom{N}{n}$.*

b) Si $r \leqslant n$ (resp. $r > n$), alors on tire au maximum r boules rouges (n boules rouges). Si $n \leqslant b$ (resp. $n > b$), alors on tire au minimum 0 (resp. $n - b$) boules rouges.
On a donc $\boxed{X(\Omega) = [\![\max(0, n-b), \min(r, n)]\!].}$

Remarque : *On pourrait aussi dire que $X(\Omega) = [\![0, n]\!]$ pour simplifier les calculs mais alors certaines valeurs seront prises par X avec probabilité nulle.*

c) Soit $k \in X(\Omega)$. L'événement $[X = k]$ est l'ensemble des tirages de k boules rouges et $n-k$ boules bleues. Pour obtenir un tel tirage, on peut d'abord choisir les tirages où on obtient une rouge (il y a $\binom{n}{k}$ façons) puis tirer les rouges (il y a r choix pour la première, $r-1$ pour la deuxième, etc. $r-k+1$ pour la dernière) et enfin tirer les bleues (il y a b choix pour la première, $b-1$ pour la deuxième, etc. $b-(n-k)+1$ pour la dernière). Ainsi

$$\begin{aligned}\mathrm{card}([X = k]) &= \binom{n}{k}\frac{r!}{(r-k)!}\,\frac{(N-r)!}{(N-r-n+k)!}\\ &= \frac{n!}{(n-k)!k!}\,\frac{r!}{(r-k)!}\,\frac{(N-r)!}{(N-r-n+k)!}\\ &= n!\binom{r}{k}\binom{N-r}{n-k}.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{P}(X = k) = \frac{\mathrm{card}([X = k])}{\mathrm{card}(\Omega)} = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \frac{\binom{pN}{k}\binom{(1-p)N}{n-k}}{\binom{N}{n}}.}$$

Remarque : *Si on prend $X(\Omega) = [\![0, n]\!]$, les formules utilisées deviennent fausses dès lors que $k \notin [\![\max(0, n-b), \min(r, n)]\!]$. Néanmoins la formule finale se prolonge à ces cas-là puisque, par convention, pour tous entiers naturels x et y tels que $y > x$, on a $\binom{x}{y} = 0$. Notons aussi que la preuve ci-dessus est plus rapide si on considère plutôt Ω l'ensemble des tirages simultanés. En effet, $\binom{pN}{k}$ est le nombre de façons de tirer k boules parmi les $r = pN$ boules rouges et $\binom{(1-p)N}{n-k}$ est le nombre de façons de tirer $n-k$ boules bleues parmi les $N - r = (1-p)N$ boules bleues.*

d) Soient $(a, b) \in \mathbb{N}^2$ et $c \in [\![0, a+b]\!]$. Si on pose $n = c$, $r = a$ et $N = a + b$ dans l'expérience aléatoire précédente, on obtient

$$1 = \sum_{k=\max(0,c-b)}^{\min(a,c)} \mathbb{P}(X = k) = \sum_{k=\max(0,c-b)}^{\min(a,c)} \frac{\binom{a}{k}\binom{b}{c-k}}{\binom{a+b}{c}}.$$

Ainsi

$$\boxed{\binom{a+b}{c} = \sum_{k=\max(0,c-b)}^{\min(a,c)} \binom{a}{k}\binom{b}{c-k} = \sum_{k=0}^{c} \binom{a}{k}\binom{b}{c-k},}$$

puisque $\binom{a}{k} = 0$ pour tout $k > a$ et $\binom{b}{c-k} = 0$ pour tout $k < c - b$.

e) On a

$$\mathbb{E}(X) = \sum_{k=\max(0,c-b)}^{\min(a,c)} k\mathbb{P}(X = k) = \sum_{k=\max(0,n-b)}^{\min(r,n)} k\frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \sum_{k=\max(0,n-b)+1}^{\min(r,n)} \frac{r\binom{r-1}{k-1}\binom{N-r}{n-k}}{\binom{N}{n}}$$

donc, en faisant le changement d'indice $j = k - 1$

$$\mathbb{E}(X) = \frac{r}{\binom{N}{n}} \sum_{j=\max(0,n-b)}^{\min(r,n)-1} \binom{r-1}{j}\binom{N-r}{n-j-1} = \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \binom{r-1}{j}\binom{N-r}{(n-1)-j}.$$

En appliquant la formule de Vandermonde avec $a = r - 1$, $b = N - r$ et $c = n - 1$ donne :

$$\sum_{j=0}^{n-1} \binom{r-1}{j}\binom{N-r}{(n-1)-j} = \binom{r-1+N-r}{n-1} = \binom{N-1}{n-1} = \frac{n}{N}\binom{N}{n}.$$

Ainsi

$$\mathbb{E}(X) = \frac{r\frac{n}{N}\binom{N}{n}}{\binom{N}{n}} = \frac{rn}{N}$$

c'est-à-dire $\boxed{\mathbb{E}(X) = np.}$