DM3 (version B) - Oral HEC 2017 (sujet E81)

1. Question de cours

a) Définition et propriétés de la loi géométrique.

Démonstration.

- On dit qu'une v.a.r. X suit **la loi géométrique** de paramètre $p \in]0,1[$ si :

  a) $\boxed{X(\Omega) = \mathbb{N}^*}$

  b) $\boxed{\forall k \in \mathbb{N}^*, \ \mathbb{P}\big([X = k]\big) = p\ (1-p)^{k-1}}$

- Rappelons l'expérience aléatoire de référence associée à la loi géométrique de paramètre p.

  × On considère une expérience qui consiste en une infinité d'épreuves de Bernoulli indépendantes et de même paramètre de succès p.

  Alors la v.a.r. donnant le rang d'apparition du premier succès obtenu lors de cette expérience suit la loi géométrique de paramètre p.

- Considérons X une v.a.r. telle que $X \hookrightarrow \mathcal{G}(p)$.

$$\boxed{\forall k \in \mathbb{N}, \ \mathbb{P}([X > k]) \ = \ (1-p)^k}$$

Cette propriété classique caractérise la loi géométrique. Plus précisément :

$$\boxed{\begin{array}{l} \times \text{ La v.a.r. } X \text{ est à valeurs entières} \\ \times \ \forall k \in \mathbb{N}, \ \mathbb{P}([X > k]) \ = \ (1-p)^k \end{array} \quad \Leftrightarrow \quad X \hookrightarrow \mathcal{G}(p)}$$

- La loi géométrique est à perte de mémoire.
  Cette propriété stipule que pour tout $k \in \mathbb{N}$ et pour tout $\ell \in \mathbb{N}$, on a :

$$\mathbb{P}_{[X>k]}([X > k + \ell]) \ = \ \frac{\mathbb{P}([X > k+\ell])}{\mathbb{P}([X > k])} = \frac{(1-p)^{k+\ell}}{(1-p)^k} = (1-p)^\ell \ = \ \mathbb{P}([X > \ell])$$

□

b) Compléter la ligne de code **Scilab** contenant des points d'interrogation pour que la fonction `geo` suivante fournisse une simulation de la loi géométrique dont le paramètre est égal à l'argument p de la fonction.

```
function x = geo(p)
    x = 1
    while rand() ???
        x = x + 1
    end
endfunction
```

Démonstration.

Commentaire

- L'énoncé demande de fournir une simulation de la loi géométrique. En réalité, c'est plutôt l'expérience consistant en la succession d'une infinité d'épreuves de Bernoulli indépendantes et de même paramètre p qui est simulée dans le programme. Soyons plus précis sur cette idée de simulation d'expérience. Simuler une expérience informatiquement c'est la coder. Généralement, les différents résultats intermédiaires d'une expérience sont traduits par des réels. Par exemple, pour simuler une expérience consistant à effectuer un lancer de pièce qui amène Pile avec probabilité p, on va écrire un programme qui renvoie 1 avec probabilité p et 0 avec probabilité $1 - p$. On a ainsi traduit Pile par 1 et Face par 0. Ce faisant, on a introduit une v.a.r. V telle que $V \hookrightarrow \mathcal{B}(p)$. Rigoureusement, c'est cette v.a.r. V que l'on simule et pas l'épreuve de Bernoulli correspondante.
- C'est d'ailleurs l'un des intérêts majeurs des variables aléatoires : elles permettent de « traduire » des résultats de l'expérience en données numériques, bien plus exploitables pour étudier l'expérience. En effet, rappelons qu'une variable aléatoire X est une fonction de Ω dans $\mathbb{R}$. Ainsi, pour chaque **issue** ω de Ω (ω est un résultat possible de l'expérience), la v.a.r. X lui associe un **nombre réel** $X(\omega)$. En résumé, les variables aléatoires permettent d'extraire de l'information de l'expérience qui est traduite sous forme numérique. Et c'est cette traduction sous forme numérique qui est exploitée pour l'étude de l'expérience et en particulier la simulation informatique.

- Le fonctionnement du programme est le suivant :
  - × on simule les résultats successifs d'épreuves de Bernoulli de paramètre p.
    Pour se fixer les idées, on introduit pour la suite de cette question la famille $(V_k)_{k\in\mathbb{N}^*}$ de v.a.r. qui vérifie : $\forall k \in \mathbb{N}^*$, $V_k \hookrightarrow \mathcal{G}(p)$.
  - × on s'arrête dès le premier succès obtenu.
  - × on utilise une variable `x` afin de compter le nombre d'épreuves effectuées jusqu'à l'obtention du premier succès. On obtient ainsi le rang d'apparition du premier succès.
    Plus formellement, cette variable de sortie `x` est la simulation d'une v.a.r. X telle que $X \hookrightarrow \mathcal{G}(p)$ (ce qui légitime, après coup, le nom de cette variable `x`).

- **Début du programme**
  La variable `x` est initialisée à 1. Cela fait référence à la $1^{\text{ère}}$ épreuve de Bernoulli qui n'a pas encore eu lieu. La variable `x` compte ainsi avec avance le nombre d'épreuves de Bernoulli qui ont lieu au cours du programme.

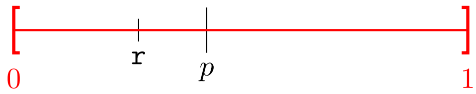
```
x = 1
```

- **Structure itérative**
  - × Il s'agit alors d'itérer tant que le succès n'est pas obtenu. On choisit donc une boucle `while`.

```
while rand() >= p
    x = x + 1
end
```

L'instruction `rand()` renvoie un réel choisi aléatoirement dans $[0, 1]$.
Plus formellement, il s'agit de simuler une v.a.r. U telle que $U \hookrightarrow \mathcal{U}([0, 1])$.

× L'instruction `rand() >= p` permet de simuler une épreuve de Bernoulli de paramètre `p`. Plus précisément, la valeur obtenue par l'appel `rand()` indique si l'épreuve de Bernoulli amène un succès ou un échec.



Considérons que l'on se trouve au $k^{\text{ème}}$ tour de boucle et que la variable `x` contient, avec une avance de 1, le nombre d'épreuves de Bernoulli qui ont eu lieu jusque là.

Deux cas se présentent.

– Si `rand()` $< p$: alors, on considère que la $k^{\text{ème}}$ épreuve de Bernoulli amène un succès. Ce cas se produit avec la probabilité attendue :

$$\mathbb{P}\left([0 \leqslant U < p]\right) = \mathbb{P}\left([U < p]\right) = p = \mathbb{P}([V_k = 1])$$

Dans ce cas, la condition de continuation de la boucle n'est pas respectée et cela met fin à l'itération. La variable `x` n'est pas mise à jour et contient alors exactement le nombre d'épreuves qui ont eu lieu. Autrement dit, en sortie de boucle, la variable `x` contient le rang d'apparition du premier succès.

– Si `rand()` $\geqslant p$: alors, on considère que la $k^{\text{ème}}$ épreuve de Bernoulli amène un échec. Ce cas se produit avec la probabilité attendue :

$$\mathbb{P}\left([p \leqslant U \leqslant 1]\right) = \mathbb{P}\left([p \leqslant U]\right) = 1 - \mathbb{P}\left([U < p]\right) = 1 - p = \mathbb{P}([V_k = 0])$$

La $k^{\text{ème}}$ épreuve de Bernoulli ayant amené un échec, on incrémente la variable `x` de 1 afin de signifier qu'une nouvelle épreuve a eu lieu. On continue alors l'itération et l'instruction `rand()` va permettre d'effectuer un nouveau tirage aléatoire dans $[0, 1]$.

Commentaire

- Afin de permettre une bonne compréhension des mécanismes en jeu, on a détaillé la réponse à cette question. Cependant, compléter correctement le programme **Scilab** démontre la bonne compréhension de la simulation demandée et permet certainement d'obtenir tous les points alloués à cette question.
- Profitons-en pour rappeler que lors de l'écriture d'un programme, on se soumet généralement à quelques règles de bonne conduite :

(1) utilisation de commentaires indiquant le but de chaque fonction,

(2) réflexion autour du découpage en sous-fonctions pouvant être réutilisées,

(3) utilisation de nom explicites pour les fonctions et le variables,

(4) indentation du code (utilisation correcte d'espaces et sauts de lignes).

Le but de ces règles est de produire un code lisible, intelligible et facilement modifiable à l'avenir. Évidemment, on ne s'attend pas, dans un sujet de concours, à ce que soit commentée la fonction dont il est demandé d'expliciter le calcul. Par contre, on s'attend à ce que les autres règles de bonne conduite soient respectées. Ne pas le faire correspond à ce que l'on nomme de l'**obfuscation** (pas forcément volontaire) de code. Sous ce terme, on désigne les méthodes permettant de rendre un code difficile à déchiffrer. Le but de telles techniques est de protéger son code. Typiquement, une entreprise ayant investi afin de développer un algorithme pourra procéder à une obfuscation de code afin que ses concurrents industriels ne puissent comprendre la manière dont procède cet algorithme.

□

Une urne contient trois jetons numérotés 1, 2 et 3.
On effectue dans cette urne, une suite de tirages d'un jeton avec remise.

2. On note Y la variable aléatoire égale au nombre de tirages nécessaires pour obtenir, pour la première fois, deux numéros successifs distincts.

a) Reconnaître la loi de la variable aléatoire $Y - 1$.

Démonstration.

- Remarquons tout d'abord : $Y(\Omega) \subset [\![2, +\infty[\![$.
  En effet :
  - $\times$ la v.a.r. Y prend des valeurs entières positives puisque Y est un nombre de tirages.
  - $\times$ la v.a.r. Y prend au minimum la valeur 2 : il est nécessaire d'attendre deux tirages pour obtenir deux numéros distincts.

$$\boxed{\text{On a : } Y(\Omega) \subset [\![2, +\infty[\![\ \text{ et ainsi } \ (Y-1)(\Omega) \subset [\![1, +\infty[\![.}$$

Commentaire

On peut démontrer : $Y(\Omega) = [\![2, +\infty[\![$. Pour ce faire, il faut démontrer : $Y(\Omega) \supset [\![2, +\infty[\![$. Cela revient à démontrer que tous les entiers plus grands que 2 sont des valeurs réellement atteintes par Y. Autrement dit, pour tout $k \geqslant 2$, on doit exhiber un ∞-tirage $\omega \in \Omega$ (l'expérience consiste à effectuer une suite infinie de tirages donc Ω est constituée d'∞-tirages) tel que : $Y(\omega) = k$. On peut par exemple choisir :

$$\omega \quad = \quad (\underbrace{1\ ,\quad 1\ ,\quad \dots,\quad 1\ ,}_{k-1 \text{ premiers tirages}}\quad 2,\quad 1\ ,\quad \dots\quad)$$

- Soit $k \in \mathbb{N}^*$.
  L'événement $[Y - 1 = k] = [Y = k+1]$ est réalisé si et seulement si il a fallu attendre le $(k+1)^{\text{ème}}$ tirage afin d'obtenir pour la première fois deux numéros distincts. Autrement dit, cet événement est réalisé si et seulement si on a obtenu le même jeton lors des k premiers tirages puis un jeton différent lors du $(k+1)^{\text{ème}}$ tirage.

  Les résultats obtenus lors des tirages ultérieurs étant libres, on peut s'intéresser uniquement aux $(k+1)$-tirages qui débutent les ∞-tirages réalisant $[Y = k+1]$.

  Un $(k+1)$-tirage qui réalise $[Y = k]$ est un $(k+1)$-uplet d'éléments de $\{1, 2, 3\}$ qui contient le même numéro dans les k premières positions et un numéro différent en $(k+1)^{\text{ème}}$ position.
  Un tel $(k+1)$-uplet est entièrement déterminé par :
  - le numéro du jeton apparaissant au k premières positions : $\binom{3}{1} = 3$ possibilités.
  - le numéro du jeton apparaissant en $(k+1)^{\text{ème}}$ position qui est différent du numéro précédent : $\binom{2}{1} = 2$ possibilités.

  Ainsi le nombre de $(k+1)$-tirages réalisant l'événement $[Y = k+1]$ est $3 \times 2 = 6$.
  Or, l'ensemble $\{1, 2, 3\}^{k+1}$ des $(k+1)$-uplets d'éléments de $\{1, 2, 3\}$ est de cardinal : 3^{k+1}.
  On en déduit :

$$\mathbb{P}\big([Y - 1 = k]\big) = \frac{\text{Card}\big([Y = k+1]\big)}{\text{Card}\big(\{1, 2, 3\}^{k+1}\big)} = \frac{2 \times 3}{3^{k+1}} = \frac{2}{3}\,\frac{3}{3^k} \;=\; \frac{2}{3}\,\left(\frac{1}{3}\right)^{k-1}$$

Finalement, on a : $\begin{cases} \times \ (Y-1)(\Omega) \subset \mathbb{N}^* \\ \times \ \forall k \in \mathbb{N}^*, \, \mathbb{P}\big([Y-1=k]\big) = \dfrac{2}{3} \left(1-\dfrac{2}{3}\right)^{k-1} \end{cases}$.

$$\boxed{\text{Ainsi : } (Y-1) \hookrightarrow \mathcal{G}\left(\frac{2}{3}\right).}$$

Commentaire

- On aurait tout aussi bien pu déterminer la loi de Y et en déduire la loi de la transformée affine $Y-1$. Pour comprendre l'intérêt de la v.a.r. $Y-1$, il est important d'insister sur ce qu'elle représente. La v.a.r. $Y-1$ prend pour valeur le nombre de tirages nécessaires entre la découverte du 1^{er} numéro apparu (celui qui apparaît lors du 1^{er} tirage) et du $2^{\text{ème}}$ non encore apparu.
- En ayant cette présentation en tête, on peut alors proposer la rédaction suivante.
  - × L'expérience consiste en la succession d'une infinité d'épreuves de Bernoulli (tirage dans l'urne) indépendantes et de même paramètre de succès $\dfrac{3-1}{3}$ (probabilité de tirer un numéro autre que celui obtenu lors du premier tirage dans l'urne contenant 3 jetons).
  - × La v.a.r. $Y-1$ est le rang d'apparition du premier succès de cette expérience (nombre de tirages nécessaires entre la découverte du 1^{er} numéro non encore apparu -ce qui se produit au premier tirage- et du $2^{\text{ème}}$).

  On en déduit : $Y-1 \hookrightarrow \mathcal{G}\left(\dfrac{2}{3}\right)$.

□

b) Déterminer l'espérance $\mathbb{E}(Y)$ et la variance $\mathbb{V}(Y)$ de la variable aléatoire Y.

Démonstration.

- Remarquons tout d'abord :
$$Y \ = \ (Y-1)+1$$
La v.a.r. Y admet une espérance (resp. une variance) car elle s'écrit comme transformée affine de la v.a.r. $Y-1$ qui admet une espérance (resp. une variance).
- De plus :
$$\begin{array}{rcll} \mathbb{E}(Y) & = & \mathbb{E}\big((Y-1)+1\big) & \\ & = & \mathbb{E}(Y-1)+\mathbb{E}(1) & \textit{(par linéarité} \\ & & & \textit{de la variance)} \\ & = & \dfrac{1}{\frac{2}{3}}+1 & \textit{(car } (Y-1) \hookrightarrow \mathcal{G}\left(\frac{2}{3}\right)\textit{)} \\ & = & \dfrac{3}{2}+1 & \\ & = & \dfrac{5}{2} & \end{array}$$

- Et :

$$\begin{array}{rcll} \mathbb{V}(Y) & = & \mathbb{V}(Y-1) & \\ & = & \dfrac{1-\frac{2}{3}}{\left(\frac{2}{3}\right)^2} & \textit{(car } (Y-1) \hookrightarrow \mathcal{G}\left(\frac{2}{3}\right)\textit{)} \\ & = & \dfrac{\frac{1}{3}}{\frac{4}{9}} & \\ & = & \dfrac{1}{3}\,\dfrac{9}{4} & \\ & = & \dfrac{3}{4} & \end{array}$$

$$\boxed{\mathbb{E}(Y) = \frac{5}{2} \text{ et } \mathbb{V}(Y) = \frac{3}{4}.}$$

□

3. On note Z la variable aléatoire égale au nombre de tirages nécessaires pour obtenir, pour la première fois, les trois numéros.

a) Soit deux entiers $k \geqslant 2$ et $\ell \geqslant 3$.
Calculer $\mathbb{P}\big([Y=k] \cap [Z=\ell]\big)$ selon les valeurs de k et ℓ.

Démonstration.

- Remarquons tout d'abord :

	L'événément $[Y=k] \cap [Z=\ell]$ est réalisé		
$\Leftrightarrow$	L'événément $[Y=k]$ est réalisé	et	l'événement $[Z=\ell]$ est réalisé
$\Leftrightarrow$	On a obtenu pour la première fois deux numéros distincts au $k^{\text{ème}}$ tirage	et	on a obtenu pour la première fois trois numéros distincts au $\ell^{\text{ème}}$ tirage

- Deux cas se présentent alors :
  × si $\ell \leqslant k$ alors :
  $$[Y=k] \cap [Z=\ell] \ = \ \varnothing$$
  En effet, il faut plus de tirages pour obtenir la première fois le $3^{\text{ème}}$ numéro non encore obtenu que pour obtenir pour la première fois le $2^{\text{ème}}$ numéro non encore obtenu.
  $$\boxed{\text{Si } \ell \leqslant k,\ \mathbb{P}\big([Y=k] \cap [Z=\ell]\big) = 0.}$$
  × si $\ell > k$, on peut raisonner comme en ***2.a)***.
  L'événément $[Y=k] \cap [Z=\ell]$ est réalisé si et seulement si il a fallu attendre le $k^{\text{ème}}$ tirage afin d'obtenir pour la première fois deux numéros distincts puis le $\ell^{\text{ème}}$ tirage pour obtenir le dernier numéro non encore obtenur.
  Autrement dit, cet événement est réalisé si et seulement si :
  - on a obtenu le même jeton lors des $k-1$ premiers tirages,
  - puis on a obtenu un autre jeton au $k^{\text{ème}}$ tirage,
  - puis on a obtenu l'un des deux numéros déjà tirés lors des tirages $k+1$ à $\ell-1$,
  - et enfin on a obtenu le dernier jeton non encore tiré lors du $\ell^{\text{ème}}$ tirage.

Les résultats obtenus lors des tirages ultérieurs étant libres, on peut s'intéresser uniquement aux ℓ-tirages qui débutent les ∞-tirages réalisant $[Y = k] \cap [Z = \ell]$.
Un ℓ-tirage qui réalise $[Y = k] \cap [Z = \ell]$ est un ℓ-uplet d'éléments de $\{1, 2, 3\}$ qui contient le même numéro dans les $k-1$ premières positions puis un autre numéro en $k^{\text{ème}}$ position puis l'un des deux numéros précédents sur les positions k à $\ell - 1$ et enfin le dernier numéro en position ℓ.
Un tel ℓ-uplet est entièrement déterminé par :

- le numéro du jeton apparaissant au $k-1$ premières positions : $\binom{3}{1} = 3$ possibilités.
- le numéro du jeton apparaissant à la $k^{\text{ème}}$ position : $\binom{2}{1} = 2$ possibilités.
- le numéro du jeton apparaissant à la $k+1^{\text{ème}}$ position : $\binom{2}{1} = 2$ possibilités.
- ...
- le numéro du jeton apparaissant à la $(\ell-1)^{\text{ème}}$ position : $\binom{2}{1} = 2$ possibilités.
- le numéro du jeton apparaissant en position ℓ : 1 possibilité.

Ainsi le nombre de ℓ-tirages réalisant l'événement $[Y = k] \cap [Z = \ell]$ est :

$$3 \times 2 \times (2 \times \ldots \times 2) \times 1 \ = \ 3 \times 2 \times 2^{(\ell-1)-(k+1)+1} \ = \ 3 \times 2 \times 2^{\ell-k-1} \ = \ 3 \times 2^{\ell-k}$$

Or, l'ensemble $\{1, 2, 3\}^\ell$ des ℓ-uplets d'éléments de $\{1, 2, 3\}$ est de cardinal : 3^ℓ.
On en déduit :

$$\mathbb{P}\big([Y = k] \cap [Z = \ell] \big) = \frac{\mathrm{Card}\big([Y = k] \cap [Z = \ell] \big)}{\mathrm{Card}\big(\{1,2,3\}^\ell\big)} = \frac{3 \times 2^{\ell-k}}{3^\ell} = \frac{3}{3^\ell}\, 2^\ell \frac{1}{2^k} = \frac{2^\ell}{3^{\ell-1}} \left(\frac{1}{2}\right)^k$$

$$\boxed{\text{Si } \ell > k,\ \mathbb{P}\big([Y = k] \cap [Z = \ell] \big) = \frac{2^\ell}{3^{\ell-1}} \left(\frac{1}{2}\right)^k.}$$

□

b) En déduire que, pour tout entier $\ell \geqslant 3$, on a : $\mathbb{P}([Z = \ell]) = \dfrac{2}{3} \left(\dfrac{2^{\ell-2} - 1}{3^{\ell-2}} \right)$.

Démonstration.
Soit $\ell \geqslant 3$.

- La famille $\Big([Y = k] \Big)_{k \in [\![2, +\infty[\![}$ est un système complet d'événements.
 D'après la formule des probabilités totales :

$$\begin{aligned}
\mathbb{P}\big([Z = \ell] \big) &= \sum_{k=2}^{+\infty} \mathbb{P}\big([Y = k] \cap [Z = \ell] \big) \\
&= \sum_{\substack{k = 2 \\ k < \ell}}^{+\infty} \mathbb{P}\big([Y = k] \cap [Z = \ell] \big) + \sum_{\substack{k = 2 \\ k \geqslant \ell}}^{+\infty} \cancel{\mathbb{P}\big([Y = k] \cap [Z = \ell] \big)} \\
&= \sum_{k=2}^{\ell-1} \mathbb{P}\big([Y = k] \cap [Z = \ell] \big) \\
&= \sum_{k=2}^{\ell-1} \frac{2^\ell}{3^{\ell-1}} \left(\frac{1}{2}\right)^k \\
&= \frac{2^\ell}{3^{\ell-1}} \sum_{k=2}^{\ell-1} \left(\frac{1}{2}\right)^k \\
&= \frac{2^\ell}{3^{\ell-1}} \ \frac{\left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^{(\ell-1)+1}}{1 - \frac{1}{2}}
\end{aligned}$$

- Enfin :

$$\frac{2^\ell}{3^{\ell-1}} \ \frac{\left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^{(\ell-1)+1}}{1-\frac{1}{2}} \ = \ 2 \times 2^\ell \ \frac{2^{-2}-2^{-\ell}}{3^{\ell-1}} \ = \ 2 \ \frac{2^{\ell-2}-1}{3^{\ell-1}}$$

$$\boxed{\forall \ell \geqslant 3, \ \mathbb{P}\big(\,[Z=\ell]\,\big) \ = \ \frac{2}{3} \ \left(\frac{2^{\ell-2}-1}{3^{\ell-2}}\right)}$$

$\square$

c) Calculer $\mathbb{E}(Z)$.

Démonstration.

- La v.a.r. Z admet une espérance si et seulement si la série $\sum\limits_{n \geqslant 3} n \ \mathbb{P}\big(\,[Z=n]\,\big)$ est absolument convergente. Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.
- Soit $N \geqslant 3$.

$$\begin{aligned}
\sum_{\ell=3}^{N} \ell \ \mathbb{P}\big(\,[Z=\ell]\,\big) \ &= \ \sum_{\ell=3}^{N} \ell \ \frac{2}{3} \ \left(\frac{2^{\ell-2}-1}{3^{\ell-2}}\right) \\
&= \ \sum_{\ell=2}^{N} \ell \ \frac{2}{3} \ \left(\frac{2^{\ell-2}-1}{3^{\ell-2}}\right) && \textit{(car le terme d'indice 2 de cette somme est nul)} \\
&= \ \left(\sum_{\ell=1}^{N} \ell \ \frac{2}{3} \ \left(\frac{2^{\ell-2}-1}{3^{\ell-2}}\right)\right) - \frac{2}{\cancel{3}} \ \frac{2^{-1}-1}{\cancel{3^{-1}}} \\
&= \ \left(\sum_{\ell=1}^{N} \ell \ \frac{2}{3} \ \left(\frac{2}{3}\right)^{\ell-2} \ - \ \sum_{\ell=1}^{N} \ell \ \frac{2}{3} \ \left(\frac{1}{3}\right)^{\ell-2}\right) - (1-2) \\
&= \ \left(\sum_{\ell=1}^{N} \ell \ \left(\frac{2}{3}\right)^{\ell-1} \ - \ 2 \sum_{\ell=1}^{N} \ell \ \left(\frac{1}{3}\right)^{\ell-1}\right) + 1
\end{aligned}$$

- On reconnaît les sommes partielles d'ordre N des séries géométriques dérivée première convergentes car de raisons respectives $\frac{2}{3} \in \,]-1,1[$ et $\frac{1}{3} \in \,]-1,1[$. De plus :

$$\sum_{\ell=1}^{N} \ell \ \left(\frac{2}{3}\right)^{\ell-1} \underset{N \to +\infty}{\longrightarrow} \sum_{\ell=1}^{+\infty} \ell \ \left(\frac{2}{3}\right)^{\ell-1} \ = \ \frac{1}{\left(1-\frac{2}{3}\right)^2} \ = \ \frac{1}{\left(\frac{1}{3}\right)^2} \ = \ 9$$

$$\text{et} \quad \sum_{\ell=1}^{N} \ell \ \left(\frac{1}{3}\right)^{\ell-1} \underset{N \to +\infty}{\longrightarrow} \sum_{\ell=1}^{+\infty} \ell \ \left(\frac{1}{3}\right)^{\ell-1} \ = \ \frac{1}{\left(1-\frac{1}{3}\right)^2} \ = \ \frac{1}{\left(\frac{2}{3}\right)^2} \ = \ \frac{9}{4}$$

- On en déduit que la v.a.r. Z admet une espérance. De plus :

$$\mathbb{E}(Z) \ = \ \sum_{\ell=3}^{+\infty} \ell \ \mathbb{P}\big(\,[Z=\ell]\,\big) \ = \ 9 - 2 \times \frac{9}{4} + 1 \ = \ 9 - \frac{9}{2} + 1 \ = \ \frac{9}{2} + 1 \ = \ \frac{11}{2}$$

$$\boxed{\mathbb{E}(Z) = \frac{11}{2}}$$

$\square$

4. D'une manière plus générale, calculer l'espérance de la variable aléatoire égale au nombre de tirages nécessaires pour obtenir, pour la première fois, tous les numéros, dans l'hypothèse où l'urne contient au départ n jetons, numérotés de 1 à n.

Démonstration.

- Notons N la v.a.r. qui prend pour valeur le nombre de tirages nécessaires pour obtenir, pour la première fois, tous les numéros. Ce nombre de tirages apparaît comme la somme :
  - × du nombre de tirages nécessaires pour obtenir le 1^{er} numéro. Ce nombre vaut 1.
  - × du nombre de tirages nécessaires entre la découverte du 1^{er} numéro apparu et du $2^{\text{ème}}$ non encore apparu.
  - × du nombre de tirages nécessaires entre la première découverte du $2^{\text{ème}}$ numéro apparu et du $3^{\text{ème}}$ numéro non encore apparu.
  - × ...
  - × du nombre de tirages nécessaires entre la première découverte du $(n-1)^{\text{ème}}$ numéro apparu et du dernier numéro non encore apparu.

  Le reste de la rédaction correspond à traduire formellement cette présentation.
- Pour tout $i \in [\![1, n]\!]$, on définit W_i la v.a.r. représentant le nombre de tirages nécessaires pour que, pour la première fois, i numéros distincts soient sortis. En particulier : $W_1 = 1$ et $W_n = N$.
- Pour tout $i \in [\![1, n]\!]$, on définit X_i la v.a.r. définie par :

$$\begin{cases} X_1 = W_1 = 1 \\ \forall i \in [\![2, n]\!], \ X_i = W_i - W_{i-1} \end{cases}$$

  Ainsi, pour tout $i \in [\![1, n]\!]$, X_i donne la valeur du nombre de tirages nécessaires entre la découverte du $(i-1)^{\text{ème}}$ numéro non encore apparu et du $i^{\text{ème}}$.
- Déterminons la loi des v.a.r. de la famille $(X_i)_{i \in [\![1,n]\!]}$. Soit $i \in [\![2, n]\!]$.
  - × L'expérience consiste en la succession d'une infinité d'épreuves de Bernoulli (tirage dans l'urne) indépendantes et de même paramètre de succès $\dfrac{n-i+1}{n}$ (probabilité de tirer un $i^{\text{ème}}$ numéro non encore obtenu dans l'urne contenant n numéro).
  - × La v.a.r. X_i est le rang d'apparition du premier succès de cette expérience (nombre de tirages nécessaires entre la découverte du $(i-1)^{\text{ème}}$ numéro non encore apparu et du $i^{\text{ème}}$).

$$\boxed{\text{On en déduit : } \forall i \in [\![2, n]\!], \ X_i \hookrightarrow \mathcal{G}\left(\frac{n-i+1}{n}\right).}$$

- Par ailleurs :

$$\begin{aligned} \sum_{i=1}^{n} X_i &= W_1 + \sum_{i=2}^{n} (W_i - W_{i-1}) \\ &= W_1 + (W_n - W_1) \qquad \textit{(par télescopage)} \\ &= W_n \\ &= N \end{aligned}$$

- La v.a.r. N admet une espérance comme somme de v.a.r. qui admettent une espérance. De plus :

$$\begin{array}{rcll}
\mathbb{E}(N) & = & \mathbb{E}\left(\sum\limits_{i=1}^{n} X_i\right) & \\
& = & \sum\limits_{i=1}^{n} \mathbb{E}(X_i) & \textit{(par linéarité de l'espérance)} \\
& = & \mathbb{E}(1) + \sum\limits_{i=2}^{n} \dfrac{1}{\frac{n-i+1}{n}} & \textit{(car pour tout } i \in [\![2,n]\!], \\
& & & X_i \hookrightarrow \mathcal{G}\left(\frac{n-i+1}{n}\right)\textit{)} \\
& = & 1 + \sum\limits_{i=2}^{n} \dfrac{n}{n-i+1} & \\
& = & \sum\limits_{i=1}^{n} \dfrac{n}{n-i+1} & \\
& = & n \sum\limits_{i=1}^{n} \dfrac{n}{n-i+1} & \\
& = & n \sum\limits_{i=1}^{n} \dfrac{1}{i} &
\end{array}$$

$$\boxed{\text{On en conclut : } \mathbb{E}(N) = n \left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right).}$$

□

Prenons un peu de recul . . .

- La dernière question de cet énoncé paraît bien sèche. Si les premières questions traitent d'un cas particulier avec n petit ($n = 3$), la présentation faite de cette question ne donne aucun indice sur la manière dont il faut procéder pour généraliser. Si le but de la question ***3.*** est uniquement de déterminer l'espérance de Z, il aurait été certainement plus pertinent :
  - × d'introduire la v.a.r. X constante égale à 1 qui représente le nombre de tirages nécessaires pour que, pour la première fois, 1 numéro distinct soit sorti (cela se produit dès le 1^{er} tirage).
  - × de déterminer la loi de $Z - Y$ en lieu et place de la loi du couple (X, Y). Cette v.a.r. prend pour valeur le nombre de tirages nécessaires entre la découverte du $2^{\text{ème}}$ numéro non encore apparu et du $3^{\text{ème}}$. Elle suit la loi géométrique de paramètre $\frac{1}{3}$.

  On peut alors déterminer l'espérance de Z en remarquant :

  $$Z = X + (Y - X) + (Z - Y)$$

  Cette écriture justifie d'ailleur l'étude de la v.a.r. $Y - 1$ $(= Y - X)$ de la question ***2.a)***.
- En optant pour la présentation décrit dans le point précédent, les questions ***3.a)*** et ***3.b)*** ne sont plus nécessaires. Pour autant, elles sont intéressantes car elles permettent de vérifier que le candidat est capable :
  - × de déterminer la loi d'un couple,
  - × d'obtenir une loi marginale à l'aide d'une loi de couple.
    Ce deuxième point ne peut être évalué que si le candidat a obtenu le résultat du point précédent. De ce fait, il aurait certainement été plus pertinent de fournir aux candidats la réponse à la ***3.a)*** plutôt que la réponse à la ***3.b)***.