Planche HEC : couples de var discrètes

Exercice avec préparation 1

1. Question de cours

a) Définition et propriétés de la loi géométrique.

b) Compléter la ligne de code **Python** 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.

```
def geo(p):
    x = 1
    while rd.random() ???
        x = x + 1
    return x
```

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 $Y - 1$.

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

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 $\mathbb{P}([Y = k] \cap [Z = \ell])$ selon les valeurs de k et ℓ.

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

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

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.

Réponses de l'exercice avec préparation 1 :

1. a) • 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, en notant $q = 1-p$:

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

• 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) On propose le code suivant :

```
def geo(p):
    x = 1
    while rd.random() > p:
        x = x + 1
    return x
```

2. a) • Remarquons tout d'abord : $Y(\Omega) \subset [\![2, +\infty[\![$.
En effet :

× la v.a.r. Y prend des valeurs entières positives puisque Y est un nombre de tirages.

× 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 :

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

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

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

b) • 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 :

$$\begin{array}{rcll} \mathbb{E}(Y) & = & \mathbb{E}\big((Y-1)+1\big) & \\ & = & \mathbb{E}(Y-1)+\mathbb{E}(1) & \textit{(par linéarité de la variance)} \\ & = & \dfrac{1}{\frac{2}{3}}+1 & \textit{(car } (Y-1)\hookrightarrow\mathcal{G}\left(\frac{2}{3}\right)\textit{)} \\ & = & \dfrac{3}{2}+1 & \\ & = & \dfrac{5}{2} & \end{array}$$

- Et :

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

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

3. a) • Remarquons tout d'abord :

	L'événement $[Y = k] \cap [Z = \ell]$ est réalisé		
$\Leftrightarrow$	L'événement $[Y = k]$ 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'événement $[Y = k] \cap [Z = \ell]$ 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 :

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

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

b) 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 :

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

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

c) • 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 :

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

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

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

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

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

4. • 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$^{\text{er}}$ numéro. Ce nombre vaut 1.

× du nombre de tirages nécessaires entre la découverte du 1$^{\text{er}}$ 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 :

$$\left\{\begin{array}{l} X_1 = W_1 = 1 \\ \forall i \in [\![2,n]\!], \ X_i = W_i - W_{i-1} \end{array}\right.$$

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 $i^{\text{ème}}$).

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

- Par ailleurs :

$$\begin{array}{rcll}
\sum\limits_{i=1}^{n} X_i & = & W_1 + \sum\limits_{i=2}^{n} (W_i - W_{i-1}) & \\
& = & W_1 + (W_n - W_1) & \textit{(par télescopage)} \\
& = & W_n & \\
& = & N &
\end{array}$$

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

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

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