ESSEC II 2020

Lorsque l'on effectue des sondages, de nombreux biais statistiques peuvent apparaître : on peut par exemple avoir considéré un échantillon non-représentatif de la population, il peut y avoir un biais dans les réponses des personnes sondées... On va s'intéresser dans ce problème à ce que l'on appelle le biais par la taille : il provient du fait que si l'on choisit une personne au hasard dans la population, celle-ci a plus de chances de faire partie d'une catégorie nombreuse de la population.
Le biais par la taille est la source de nombreux « paradoxes » probabilistes, comme le fait que les gagnants du loto vivent en moyenne plus longtemps (parce que les gagnants sont ceux qui ont pu jouer au loto plus longtemps) ou le fait que vos amis ont en moyenne plus d'amis que vous (car les gens qui ont un très grand nombre d'amis font sûrement partie de vos amis). On verra ici comment formaliser le biais par la taille, et l'utiliser dans différents contextes.
Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. Pour toute variable aléatoire X, on notera $\mathbb{E}(X)$ son espérance (resp. $\mathbb{V}(X)$ sa variance) lorsqu'elles existent.

Première partie : Biais par la taille, exemples discrets

1. On suppose que le nombre d'enfants dans une famille française est une variable aléatoire X. Pour connaître la loi de X, une idée serait d'interroger les élèves d'une école pour connaître le nombre d'enfants dans leur famille.
On va voir que cette approche introduit un biais, en considérant une situation particulière. Supposons que X suive la loi binomiale de paramètres $n = 10$ et $p = \frac{1}{5}$. On note $p_k = \mathbb{P}(\,[X = k]\,)$ pour $k \in \{0, 1, \ldots, 10\}$.

a) (i) Rappeler l'expression de p_k pour $k \in \{0, 1, \ldots, 10\}$.

Démonstration.

$$\boxed{\forall k \in [\![0, 10]\!],\ p_k = \mathbb{P}([X = k]) = \binom{10}{k} \left(\frac{1}{5}\right)^k \left(1 - \frac{1}{5}\right)^{10-k}}$$

□

(ii) Que vaut $\mathbb{E}(X)$?

Démonstration.

$$\boxed{\mathbb{E}(X) \ = \ 10\ \frac{1}{5} \ = \ 2}$$

□

(iii) Donner $\mathbb{V}(X)$, et en déduire $\mathbb{E}(X^2)$.

Démonstration.

$$\boxed{\mathbb{V}(X) \ = \ 10\ \frac{1}{5}\ \frac{4}{5} \ = \ \frac{8}{5}}$$

De plus, par formule de Koenig-Huygens : $\mathbb{V}(X) \ = \ \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2$. D'où :

$$\mathbb{E}(X^2) \ = \ \mathbb{V}(X) + \big(\mathbb{E}(X)\big)^2 \ = \ \frac{8}{5} + 2^2 \ = \ \frac{8}{5} + 4$$

$$\boxed{\text{On en déduit : } \mathbb{E}(X^2) = \frac{28}{5}.}$$

□

b) Soit M_k le nombre de familles à k enfants, $M = \sum\limits_{k=0}^{10} M_k$ le nombre total de familles (donc $p_k = \dfrac{M_k}{M}$). Soit N_k le nombre total d'enfants (c'est-à-dire dans toute la population) qui font partie d'une famille à k enfants, et $N = \sum\limits_{k=0}^{10} N_k$ le nombre total d'enfants de la population.

Commentaire

- L'énoncé se permet ici une confusion entre probabilité et proportion. Rappelons la différence entre ces deux termes.
  - × Une probabilité est une valeur **théorique**. Elle provient de la modélisation d'une expérience. Ici, on modélise le nombre d'enfants dans une famille française par une v.a.r. X de loi $\mathcal{B}\left(10, \frac{1}{5}\right)$.
  - × Une proportion est une valeur **empirique**. C'est une fréquence obtenue à l'aide d'observations réelles. Ici :
$$\frac{M_k}{M} = \frac{\text{nombre de familles à } k \text{ enfants}}{\text{nombre total de familles}}$$
- Il est cependant naturel d'approcher $p_k = \mathbb{P}([X = k])$ par $\dfrac{M_k}{M}$. Pour cela l'idée est :
  - × de simuler un grand nombre de fois (M est ce grand nombre) la v.a.r. X. Formellement, on souhaite obtenir un M-uplet $(x_1, \ldots, x_M)$ qui correspond à l'observation d'un M-échantillon $(X_1, \ldots, X_M)$ de la v.a.r. X.
  - × de compter le nombre de fois où $x_i = k$.

  Cette idée est justifiée par la loi faible des grands nombres (LfGN) qui affirme :
$$\frac{M_k}{M} = \frac{\text{nombre de fois où } x_i = k}{\text{taille } (N) \text{ de l'observation}} \approx \mathbb{P}([X = k]) = p_k$$

(i) Démontrer : $N_k = k\, p_k\, M$.

Démonstration.
Il y a M_k familles de k enfants. On en déduit que le nombre d'enfants qui font partie d'une famille à k enfants est :
$$N_k = k\, M_k$$

$$\boxed{\text{Or } p_k = \frac{M_k}{M}. \text{ D'où : } N_k = k\, p_k\, M.}$$ □

(ii) Démontrer : $\dfrac{N}{M} = 2$.

Démonstration.
On calcule :
$$\begin{aligned}
\frac{N}{M} = \frac{1}{M} \sum_{k=0}^{10} N_k &= \frac{1}{M} \sum_{k=0}^{10} k\, p_k\, M && \textit{(d'après la question précédente)} \\
&= \frac{1}{\cancel{M}}\, \cancel{M} \sum_{k=0}^{10} k\, p_k \\
&= \sum_{k=0}^{10} k\, \mathbb{P}([X = k]) && \textit{(par définition de } p_k\textit{)} \\
&= \mathbb{E}(X)
\end{aligned}$$

$$\boxed{\text{D'après } \textbf{\textit{1.a)(i)}}, \text{ on obtient : } \frac{N}{M} = 2.}$$ □

(iii) Montrer que la proportion des enfants provenant d'une famille à k enfants est : $p_k^* = \dfrac{k\, p_k}{2}$.

Démonstration.
La proportion d'enfants provenant d'une famille à k enfants est :

$$p_k^* = \frac{\text{nombre d'enfants qui font partie d'une famille à } k \text{ enfants}}{\text{nombre total d'enfants}} = \frac{N_k}{N} = \frac{\frac{N_k}{M}}{\frac{N}{M}}$$

D'après les questions ***1.b)(i)*** et ***1.b)(ii)***, on obtient :

$$p_k^* = \frac{\frac{k\, p_k\, \cancel{M}}{\cancel{M}}}{2}$$

$$\boxed{\text{Finalement : } p_k^* = \frac{k\, p_k}{2}.}$$

□

c) On choisit une personne au hasard dans la rue, à qui l'on demande combien d'enfants ses parents ont eu (lui ou elle inclus). On note Y ce nombre d'enfants.

(i) Pour tout entier k élément de $\{1, 2, \dots, 10\}$, démontrer : $\mathbb{P}(\,[Y = k]\,) = \dfrac{k\, p_k}{2}$.

Démonstration.
Soit $k \in [\![1, 10]\!]$.
L'événement $[Y = k]$ est réalisé si et seulement si la personne interrogée fait partie d'une fratrie de k enfants.
En confondant proportion et probabilité (cf remarque faite en question ***1.b)***), on obtient :

$$\mathbb{P}([Y = k]) = p_k^*$$

$$\boxed{\text{D'après la question précédente, on en déduit : } \forall k \in [\![1, 10]\!],\ \mathbb{P}([Y = k]) = \frac{k\, p_k}{2}.}$$

□

(ii) Démontrer : $\mathbb{E}(Y) = \dfrac{\mathbb{E}(X^2)}{\mathbb{E}(X)}$.

Démonstration.

- Tout d'abord, par définition de la v.a.r. Y : $Y(\Omega) \subset [\![1, 10]\!]$. En effet, une personne donnée fait forcément partie d'une fratrie d'au moins 1 enfant.
  La v.a.r. Y admet alors une espérance, car c'est une v.a.r. finie.
- Par définition de l'espérance :

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{k=1}^{10} k\, \mathbb{P}([Y = k]) \\
&= \sum_{k=1}^{10} k\, \frac{k\, p_k}{2} && \textit{(d'après la question précédente)} \\
&= \frac{1}{2} \sum_{k=1}^{10} k^2\, \mathbb{P}([X = k]) \\
&= \frac{1}{2}\, \mathbb{E}(X^2) && \textit{(par définition du moment d'ordre 2)}
\end{aligned}$$

$$\boxed{\text{Ainsi, d'après la question } \textbf{\textit{1.a)(ii)}} : \mathbb{E}(Y) = \frac{\mathbb{E}(X^2)}{\mathbb{E}(X)}.}$$

□

(iii) En déduire $\mathbb{E}(Y)$ et le comparer à $\mathbb{E}(X)$.

Démonstration.

- D'après les questions ***1.a)(i)*** et ***1.a)(ii)*** :

$$\mathbb{E}(Y) = \frac{\mathbb{E}(X^2)}{\mathbb{E}(X)} = \frac{\frac{28}{5}}{2} = \frac{1}{2}\,\frac{28}{5} = \frac{14}{5}$$

$$\boxed{\mathbb{E}(Y) = \frac{14}{5}}$$

- On remarque :

$$\underset{\mathbb{E}(Y)}{\underset{\shortparallel}{\frac{14}{5}}} \geqslant \frac{10}{5} = \underset{\mathbb{E}(X)}{\underset{\shortparallel}{2}}$$

$$\boxed{\text{Ainsi : } \mathbb{E}(Y) \geqslant \mathbb{E}(X).}$$

□

Commentaire

L'énoncé nous a présenté dans cette question ***1.*** un exemple de biais par la taille. En effet :

- on cherche à connaître la loi de X, nombre d'enfants dans une famille française,
- on interroge pour cela des enfants en leur demandant la taille de leur fratrie. On modélise cette taille par une v.a.r. Y.

Dans cette expérience, l'échantillon considéré n'est pas représentatif des familles françaises. En effet, en demandant aux enfants la taille de leur fratrie, il est impossible de prendre en compte les familles sans enfant. On a d'ailleurs précisé dans les questions précédentes : $X(\Omega) = [\![0, 10]\!]$ mais $Y(\Omega) \subset [\![1, 10]\!]$).

C'est cette non-représentativité qui induit un biais dans l'étude de la loi de X.

2. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$, non identiquement nulle et admettant une espérance. Pour tout entier $i > 0$, on pose : $q_i = \dfrac{i}{\mathbb{E}(X)}\, \mathbb{P}(\,[X = i]\,)$.

Commentaire

- Remarquons que l'hypothèse émise par l'énoncé « la v.a.r. X est non identiquement nulle » ne suffit pas pour que, pour tout $i \in \mathbb{N}^*$, le réel q_i soit bien défini, *i.e.* $\mathbb{E}(X) \neq 0$. En effet, si on définit une v.a.r. X par :

$$X(\Omega) = \{0, 1\}, \qquad \mathbb{P}([X = 0]) = 1 \quad \text{et} \quad \mathbb{P}([X = 1]) = 0$$

  Cette variable aléatoire :

  × est à valeurs dans $\mathbb{N}$ (car $\{0, 1\} \subset \mathbb{N}$), n'est pas identiquement nulle car $1 \in X(\Omega)$,

  × admet une espérance car c'est une v.a.r. finie et :

$$\mathbb{E}(X) = 0 \times \mathbb{P}([X = 0]) + 1 \times \mathbb{P}([X = 1]) = 0$$

- La bonne hypothèse est la suivante : « la v.a.r. X n'est pas nulle presque sûrement » (*i.e.* $\mathbb{P}([X = 0]) < 1$). Dans ce cas :

  × comme X est une v.a.r. à valeurs positives ($X(\Omega) \subset \mathbb{N}$), alors, par positivité de l'espérance : $\mathbb{E}(X) \geqslant 0$.

  × Démontrons par l'absurde : $\mathbb{E}(X) > 0$.
  Supposons : $\mathbb{E}(X) \leqslant 0$. D'après le point précédent, on en déduit : $\mathbb{E}(X) = 0$.
  De plus la v.a.r. X est à valeurs positives. On en conclut que X est nulle presque sûrement ($\mathbb{P}([X = 0]) = 1$). Absurde !

  On obtient bien : $\mathbb{E}(X) > 0$. Ainsi, pour tout $i \in \mathbb{N}^*$, q_i est bien défini.

a) Calculer $\sum\limits_{i=1}^{+\infty} q_i$.

Démonstration.

- Soit $N \in \mathbb{N}$.

$$\sum_{i=1}^{N} q_i \;=\; \sum_{i=1}^{N} \frac{i}{\mathbb{E}(X)}\,\mathbb{P}([X=i]) \;=\; \frac{1}{\mathbb{E}(X)} \sum_{i=1}^{N} i\,\mathbb{P}([X=i])$$

- De plus, la v.a.r. X admet une espérance. On en déduit que la série $\sum\limits_{i\geqslant 1} i\,\mathbb{P}([X=i])$ est absolument convergente, donc convergente. Ainsi la série $\sum\limits_{i\geqslant 1} q_i$ est convergente et :

$$\sum_{i=1}^{+\infty} q_i \;=\; \frac{1}{\mathbb{E}(X)} \sum_{i=1}^{+\infty} i\,\mathbb{P}([X=i]) \;=\; \frac{1}{\mathbb{E}(X)}\,\mathbb{E}(X) \;=\; 1$$

$$\boxed{\sum_{i=1}^{+\infty} q_i = 1}$$

□

La suite $(q_i)_{i>0}$ définie ci-dessus définit donc bien une loi de probabilité.

Commentaire

Rappelons que pour définir une loi de probabilité, la suite $(q_i)_{i\in\mathbb{N}^*}$ doit vérifier :
1) $\forall i \in \mathbb{N}^*$, $q_i \geqslant 0$,
2) $\sum\limits_{i\geqslant 1} q_i$ est convergente et sa somme vaut 1.
Le point ***2)*** a été démontré en question précédente. Justifions le point ***1)***. Soit $i \in \mathbb{N}^*$.

- Comme $\mathbb{P}$ est une application probabilité : $\mathbb{P}([X=i]) \geqslant 0$.
- De plus la v.a.r. X est à valeurs positives ($X(\Omega) \subset \mathbb{N}$). Par positivité de l'espérance : $\mathbb{E}(X) \geqslant 0$.

Enfin : $i \geqslant 0$. D'où : $q_i \geqslant 0$.

On considère la variable aléatoire X^* dont la loi est donnée par les q_i, c'est-à-dire, pour tout i entier naturel non nul :

$$\mathbb{P}(\,[X^*=i]\,) \;=\; \frac{i}{\mathbb{E}(X)}\,\mathbb{P}(\,[X=i]\,)$$

On dit que X^* suit *la loi de X biaisée par la taille.*

b) On suppose que X admet un moment d'ordre 2. Démontrer : $\mathbb{E}(X^*) = \dfrac{\mathbb{E}(X^2)}{\mathbb{E}(X)}$

Démonstration.

- La v.a.r. X^* admet une espérance si et seulement si la série $\sum\limits_{i\geqslant 1} i\,\mathbb{P}([X^*=i])$ est absolument convergente. Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.
- Soit $N \in \mathbb{N}$.

$$\sum_{i=0}^{N} i\,\mathbb{P}([X^*=i]) \;=\; \sum_{i=0}^{N} i\,\frac{i}{\mathbb{E}(X)}\,\mathbb{P}([X=i]) \;=\; \frac{1}{\mathbb{E}(X)} \sum_{i=0}^{N} i^2\,\mathbb{P}([X=i])$$

- La v.a.r. X admet un moment d'ordre 2. On en déduit que la série $\sum\limits_{i\geqslant 0} i^2\,\mathbb{P}([X=i])$ est absolument convergente, et donc convergente. Ainsi, la v.a.r. X^* admet une espérance et :

$$\mathbb{E}(X^*) \ = \ \frac{1}{\mathbb{E}(X)} \sum_{i=0}^{+\infty} i^2\,\mathbb{P}([X=i]) \ = \ \frac{\mathbb{E}(X^2)}{\mathbb{E}(X)}$$

$$\boxed{\mathbb{E}(X^*) = \frac{\mathbb{E}(X^2)}{\mathbb{E}(X)}}$$

□

c) En déduire que si $\mathbb{E}(X^2)$ existe, on a : $\mathbb{V}(X) = \mathbb{E}(X)\,\big(\mathbb{E}(X^*) - \mathbb{E}(X)\big)$.

Démonstration.
Supposons que la v.a.r. X admet un moment d'ordre 2. Alors X admet une variance. De plus :

$$\begin{array}{rcll} \mathbb{V}(X) & = & \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2 & \textit{(par formule de Koenig-Huygens)} \\ & = & \mathbb{E}(X)\,\mathbb{E}(X^*) - \big(\mathbb{E}(X)\big)^2 & \textit{(d'après la question précédente)} \end{array}$$

$$\boxed{\text{Si } \mathbb{E}(X^2) \text{ existe, alors : } \mathbb{V}(X) = \mathbb{E}(X)\,\big(\mathbb{E}(X^*) - \mathbb{E}(X)\big).}$$

□

d) Conclure : $\mathbb{E}(X^*) \geqslant \mathbb{E}(X)$.

Démonstration.
Supposons que X admet un moment d'ordre 2. Alors :

- les v.a.r. X et X^* admettent une espérance (d'après ***2.b)***).
- d'après ***2.c)*** :
$$\mathbb{V}(X) \ = \ \mathbb{E}(X)\,\big(\mathbb{E}(X^*) - \mathbb{E}(X)\big)$$
Or une variance est toujours positive. D'où : $\mathbb{E}(X)\,\big(\mathbb{E}(X^*) - \mathbb{E}(X)\big) \geqslant 0$.
- Comme la v.a.r. X est à valeurs positives, par positivité de l'espérance : $\mathbb{E}(X) \geqslant 0$.
De plus : $\mathbb{E}(X) \neq 0$ (démontré dans la remarque en début de question ***2.***).
Ainsi : $\mathbb{E}(X) > 0$. Donc $\mathbb{V}(X)$ est du signe de $\mathbb{E}(X^*) - \mathbb{E}(X)$. On en déduit :
$$\mathbb{E}(X^*) - \mathbb{E}(X) \ \geqslant \ 0$$

$$\boxed{\text{Finalement : } \mathbb{E}(X^*) \geqslant \mathbb{E}(X).}$$

Commentaire

Notons que la **stricte** positivité de $\mathbb{E}(X)$ est indispensable pour conclure.
En effet, si $\mathbb{E}(X) = 0$, alors :

$$\mathbb{V}(X) \ = \ \mathbb{E}(X)\,\big(\mathbb{E}(X^*) - \mathbb{E}(X)\big) \ = \ 0$$

Et dans ce cas, le signe de $\mathbb{E}(X^*) - \mathbb{E}(X)$ peut tout aussi bien être positif ou négatif.

□

3. a) Soit λ un réel strictement positif. On suppose que X est une variable aléatoire qui suit la loi de Poisson de paramètre λ. Soit X^* une variable aléatoire suivant la loi de X biaisée par la taille.

(i) Donner la loi de X^*.

Démonstration.

- En utilisant la définition de X^* (qui est bien définie car X est à valeurs dans $\mathbb{N}$, n'est pas nulle presque sûrement et admet une espérance), on considère dans la suite : $X^*(\Omega) = \mathbb{N}^*$.

- Soit $i \in \mathbb{N}^*$.

$$\begin{array}{rcll} \mathbb{P}([X^* = i]) & = & \dfrac{i}{\mathbb{E}(X)}\ \mathbb{P}([X = i]) & \textit{(par définition de } X^*\textit{)} \\ & = & \dfrac{i}{\lambda}\ \mathrm{e}^{-\lambda}\ \dfrac{\lambda^i}{i!} & \textit{(car } X \hookrightarrow \mathcal{P}(\lambda)\textit{)} \end{array}$$

$$\boxed{\begin{array}{c}\text{Finalement : } X^*(\Omega) = \mathbb{N}^* \text{ et} \\ \forall i \in \mathbb{N}^*,\ \mathbb{P}([X^* = i]) = \mathrm{e}^{-\lambda}\ \dfrac{\lambda^{i-1}}{(i-1)!}.\end{array}}$$

□

Commentaire

- Profitons de cette question pour faire un point sur la notation $X(\Omega)$.
  Rappelons qu'une v.a.r. X est une application $X : \Omega \to \mathbb{R}$.
  Comme la notation le suggère, $X(\Omega)$ est l'image de Ω par l'application X.
  Ainsi, $X(\Omega)$ n'est rien d'autre que l'ensemble des valeurs prises par la v.a.r. X :

$$\begin{array}{rcl} X(\Omega) & = & \{X(\omega) \mid \omega \in \Omega\} \\ & = & \{x \in \mathbb{R} \mid \exists \omega \in \Omega,\ X(\omega) = x\} \end{array}$$

  Il faut bien noter que dans cette définition aucune application probabilité $\mathbb{P}$ n'apparaît.
- Il est toujours correct d'écrire : $X(\Omega) \ \subseteq\ ]-\infty, +\infty[$.
  En effet, cette propriété signifie que toute v.a.r. X est à valeurs dans $\mathbb{R}$, ce qui est toujours le cas par définition de la notion de variable aléatoire.
- Dans le cas des v.a.r. **discrètes**, il est d'usage relativement courant de confondre :
  × l'ensemble de valeurs possibles de la v.a.r. X (*i.e.* l'ensemble $X(\Omega)$),
  × l'ensemble $\{x \in \mathbb{R} \mid \mathbb{P}\big([X = x]\big) \neq 0\}$, ensemble des valeurs que X prend avec probabilité non nulle. Dans le cas qui nous intéresse ici, à savoir X est une v.a.r. discrète, cet ensemble est appelé support de X et est noté $\mathrm{Supp}(X)$.
- Si X est une v.a.r. **discrète**, il est à noter que toute valeur prise par X avec probabilité non nulle est une valeur prise par X. Autrement dit, on a toujours :

$$\boxed{\mathrm{Supp}(X) \subseteq X(\Omega)}$$

  En effet, si $x \in \mathrm{Supp}(X)$ alors $\mathbb{P}\big([X = x]\big) \neq 0$. On en déduit : $[X = x] \neq \varnothing$. Il existe donc (au moins) un élément $\omega \in \Omega$ tel que $X(\omega) = x$. La v.a.r. X prend donc la valeur x.
- La réciproque n'est pas forcément vérifiée : $X(\Omega) \cancel{\subseteq} \mathrm{Supp}(X)$.
  Autrement dit, une v.a.r. X peut prendre une valeur avec probabilité nulle. On peut par exemple penser à l'expérience consistant au lancer d'un dé à 6 faces. La v.a.r. X qui donne le résultat du dé a pour ensemble image $X(\Omega) = [\![1, 6]\!]$. Si on considère que le dé est truqué et ne renvoie que 6, alors le support de X est $\mathrm{Supp}(X) = \{6\}$.
- Ici, seul le support de X^* est précisé et non son ensemble image. Pour l'étude de la v.a.r. X, on se permet alors d'écrire :

  « Dans la suite, on **considère** : $X^*(\Omega) = \mathbb{N}^*$ $(= \mathrm{Supp}(X))$ »

  En **décrétant** la valeur de $X^*(\Omega)$, on ne commet pas une erreur mais on décide d'ajouter une hypothèse qui ne fait pas partie de l'énoncé. Cette audace permet de travailler avec un ensemble image connu, ce qui permet de structurer certaines démonstrations.

(ii) Vérifier que X^* suit la même loi que $X+1$.

Démonstration.

- Tout d'abord, comme $X \hookrightarrow \mathcal{P}(\lambda)$, alors : $X(\Omega) = \mathbb{N}$. En notant $Y = X+1$, on en déduit : $Y(\Omega) = \mathbb{N}^*$.
- Soit $i \in \mathbb{N}^*$.

$$\mathbb{P}([Y=i]) \;=\; \mathbb{P}([X+1=i]) \;=\; \mathbb{P}([X=i-1]) \;=\; \mathrm{e}^{-\lambda}\,\frac{\lambda^{i-1}}{(i-1)!} \qquad \textit{(car } i-1 \in \mathbb{N}\textit{)}$$

- Avec la question précédente, on remarque :
  - × $X^*(\Omega) = \mathbb{N}^* = (X+1)(\Omega)$
  - × $\forall i \in \mathbb{N}^*$, $\mathbb{P}([X^*=i]) = \mathbb{P}([X+1=i])$

$$\boxed{\text{On en conclut que } X^* \text{ et } X+1 \text{ ont même loi.}}$$

□

b) Réciproquement, on suppose que X est une variable aléatoire à valeurs dans $\mathbb{N}$ admettant une espérance non nulle, telle que X^* et $X+1$ suivent la même loi.

(i) Montrer que pour tout $k \geqslant 1$: $\mathbb{P}(\,[X=k]\,) = \dfrac{\mathbb{E}(X)}{k}\,\mathbb{P}(\,[X=k-1]\,)$.

Démonstration.
Soit $k \in \mathbb{N}^*$. Comme X^* et $X+1$ on même loi :

$$\begin{array}{rrcll} & \mathbb{P}([X^*=k]) & = & \mathbb{P}([X+1=k]) & \\ \text{donc} & \dfrac{k}{\mathbb{E}(X)}\,\mathbb{P}([X=k]) & = & \mathbb{P}([X=k-1]) & \textit{(par définition de } X^*\textit{)} \end{array}$$

$$\boxed{\text{On en déduit : } \forall k \in \mathbb{N}^*,\ \mathbb{P}([X=k]) = \frac{\mathbb{E}(X)}{k}\,\mathbb{P}([X=k-1]).}$$

□

(ii) Montrer que pour tout k entier naturel : $\mathbb{P}(\,[X=k]\,) = \dfrac{\big(\mathbb{E}(X)\big)^k}{k!}\,\mathbb{P}(\,[X=0]\,)$.

Démonstration.

Démontrons par récurrence : $\forall k \in \mathbb{N}, \mathcal{P}(k)$ où $\mathcal{P}(k) : \mathbb{P}([X=k]) = \dfrac{\big(\mathbb{E}(X)\big)^k}{k!}\,\mathbb{P}([X=0])$.

▸ **Initialisation**
On remarque :

$$\frac{\big(\mathbb{E}(X)\big)^0}{0!}\,\mathbb{P}([X=0]) \;=\; \frac{1}{1}\,\mathbb{P}([X=0]) \;=\; \mathbb{P}([X=0])$$

D'où $\mathcal{P}(0)$.

▸ **Hérédité** : soit $k \in \mathbb{N}$.
Supposons $\mathcal{P}(k)$ et démontrons $\mathcal{P}(k+1)$ (*i.e.* $\mathbb{P}([X=k+1]) = \dfrac{\big(\mathbb{E}(X)\big)^{k+1}}{(k+1)!}\,\mathbb{P}([X=0])$).
Comme $k+1 \in \mathbb{N}^*$, d'après la question précédente :

$$\mathbb{P}([X=k+1]) \;=\; \frac{\mathbb{E}(X)}{k+1}\,\mathbb{P}([X=(k+1)-1]) \;=\; \frac{\mathbb{E}(X)}{k+1}\,\mathbb{P}([X=k])$$

Par hypothèse de récurrence, on obtient :

$$\mathbb{P}([X = k+1]) \;=\; \frac{\mathbb{E}(X)}{k+1} \, \frac{\big(\mathbb{E}(X)\big)^k}{k!} \, \mathbb{P}([X=0]) \;=\; \frac{\big(\mathbb{E}(X)\big)^{k+1}}{(k+1)!} \, \mathbb{P}([X=0])$$

D'où $\mathcal{P}(k+1)$.

$$\boxed{\text{Par principe de récurrence : } \forall k \in \mathbb{N},\ \mathbb{P}([X=k]) = \frac{\big(\mathbb{E}(X)\big)^k}{k!} \, \mathbb{P}([X=0]).}$$

□

(iii) En déduire la loi de X.

Démonstration.

- On sait déjà d'après l'énoncé : $X(\Omega) \subset \mathbb{N}$.
- D'après la question précédente : $\forall k \in \mathbb{N}$, $\mathbb{P}([X=k]) = \dfrac{\big(\mathbb{E}(X)\big)^k}{k!} \, \mathbb{P}([X=0])$. Il reste donc à déterminer la valeur de $\mathbb{P}([X=0])$.
  × La famille $\big([X=k] \big)_{k\in\mathbb{N}}$ forme un système complet d'événements. On en déduit :
  $$\sum_{k=0}^{+\infty} \mathbb{P}([X=k]) \;=\; 1$$
  × Or :
  $$\begin{aligned} \sum_{k=0}^{+\infty} \mathbb{P}([X=k]) &= \sum_{k=0}^{+\infty} \frac{\big(\mathbb{E}(X)\big)^k}{k!} \, \mathbb{P}([X=0]) && \textit{(d'après la question précédente)} \\ &= \mathbb{P}([X=0]) \sum_{k=0}^{+\infty} \frac{\big(\mathbb{E}(X)\big)^k}{k!} \\ &= \mathbb{P}([X=0]) \, \mathrm{e}^{\mathbb{E}(X)} && \textit{(car } \sum_{k\geqslant 0} \frac{\big(\mathbb{E}(X)\big)^k}{k!} \textit{ est une série exponentielle de paramètre } \mathbb{E}(X)\textit{)} \end{aligned}$$
  On en déduit : $\mathbb{P}([X=0]) \, \mathrm{e}^{\mathbb{E}(X)} = 1$.
  $$\boxed{\text{D'où : } \mathbb{P}([X=0]) = \mathrm{e}^{-\mathbb{E}(X)}.}$$
  × Ainsi, pour tout $k \in \mathbb{N}$:
  $$\mathbb{P}([X=k]) \;=\; \frac{\big(\mathbb{E}(X)\big)^k}{k!} \, \mathrm{e}^{-\mathbb{E}(X)}$$

$$\boxed{\begin{array}{c}\text{De plus : } \mathbb{E}(X) > 0 \text{ (car } X \text{ est à valeurs positives et d'espérance non nulle).} \\ \text{On en conclut : } X \hookrightarrow \mathcal{P}\left(\mathbb{E}(X)\right).\end{array}}$$

□

Commentaire

Dans cette question ***3.***, on a démontré :

$$X \hookrightarrow \mathcal{P}(\lambda) \quad \underset{\textit{3.b)}}{\overset{\textit{3.a)}}{\rightleftarrows}} \quad \left\{ \begin{array}{l} X(\Omega) \subset \mathbb{N} \\ X \text{ admet une espérance } \lambda > 0 \\ X^* \text{ et } X+1 \text{ ont même loi} \end{array} \right.$$

On a ainsi démontré que les seules v.a.r. X à valeurs entières telles que X^* et $X+1$ ont même loi sont celles qui suivent une loi de Poisson.

4. Le paradoxe du temps d'attente du bus.
Soit $n \geqslant 1$ un entier naturel, et soit X une variable aléatoire à valeurs dans $\{1, \dots, n\}$ telle que pour tout $1 \leqslant k \leqslant n$: $\mathbb{P}(\,[X = k]\,) > 0$. On suppose qu'à un arrêt de bus donné, les intervalles de temps entre deux bus consécutifs, exprimés en minutes, sont des variables aléatoires indépendantes, de même loi que X. Une personne arrive à cet arrêt à un instant aléatoire, et se demande combien de temps elle va attendre.

a) Une première idée est que la personne arrive à un instant uniforme entre deux arrivées de bus, séparées par un intervalle de X minutes. On note T la variable aléatoire qui représente le temps d'attente (à valeurs dans $[\![1, n]\!]$) et on suppose donc que pour tout entier k élément de $[\![1, n]\!]$:

$$\mathbb{P}_{[X=k]}(\,[T = j]\,) = \begin{cases} \dfrac{1}{k} & \text{si } j \in \{1, \dots, k\} \\ 0 & \text{si } j > k \end{cases}$$

(i) Montrer que pour tout entier $k \in \{1, \dots, n\}$, on a : $\displaystyle\sum_{j=1}^{n} j\, \mathbb{P}_{[X=k]}(\,[T = j]\,) = \frac{k+1}{2}$.

Démonstration.
Soit $k \in [\![1, n]\!]$.

$$\begin{aligned} \sum_{j=1}^{n} j\ \mathbb{P}_{[X=k]}([T = j]) &= \sum_{\substack{j=1 \\ j \in [\![1,k]\!]}}^{n} \mathbb{P}_{[X=k]}(\,[T = j]\,) + \sum_{\substack{j=1 \\ j > k}}^{n} \mathbb{P}_{[X=k]}(\,[T = j]\,) \\ &= \sum_{j=1}^{k} j\ \mathbb{P}_{[X=k]}([T = j]) + \cancel{\sum_{j=k+1}^{n} j\ \mathbb{P}_{[X=k]}([T = j])} \\ &= \sum_{j=1}^{k} j\ \frac{1}{k} \ = \ \frac{1}{k} \sum_{j=1}^{k} j \\ &= \frac{1}{\cancel{k}}\ \frac{\cancel{k}\ (k+1)}{2} \end{aligned}$$

$$\boxed{\text{On en déduit : } \forall k \in [\![1, n]\!],\ \sum_{j=1}^{n} j\ \mathbb{P}_{[X=k]}([T = j]) = \frac{k+1}{2}.}$$

□

(ii) En déduire : $\displaystyle\sum_{k=1}^{n} \sum_{j=1}^{n} j\, \mathbb{P}(\,[X = k]\,)\, \mathbb{P}_{[X=k]}(\,[T = j]\,) = \frac{\mathbb{E}(X+1)}{2}$.

Démonstration.

- La v.a.r. X est finie. Elle admet donc une espérance. On en déduit que la v.a.r. $X + 1$ admet une espérance en tant que transformée affine de X qui en admet une.
- De plus : $\displaystyle\sum_{k=1}^{n} \left(\sum_{j=1}^{n} j\ \mathbb{P}([X = k]\ \ \mathbb{P}_{[X=k]}([T = j]) \right)$

$$\begin{aligned} &= \sum_{k=1}^{n} \mathbb{P}([X = k]) \left(\sum_{j=1}^{n} j\ \mathbb{P}_{[X=k]}([T = j]) \right) \\ &= \sum_{k=1}^{n} \mathbb{P}([X = k])\ \frac{k+1}{2} && \textit{(d'après la question précédente)} \\ &= \frac{1}{2} \sum_{k=1}^{n} (k+1)\ \mathbb{P}([X = k]) \\ &= \frac{1}{2}\ \mathbb{E}(X+1) && \textit{(par théorème de transfert)} \end{aligned}$$

$$\boxed{\sum_{k=1}^{n} \left(\sum_{j=1}^{n} j\ \mathbb{P}([X = k])\ \mathbb{P}_{[X=k]}([T = j]) \right) = \frac{\mathbb{E}(X+1)}{2}}$$

□

(iii) Démontrer : $\mathbb{E}(T) = \sum\limits_{j=1}^{n} \sum\limits_{k=1}^{n} j\, \mathbb{P}(\, [X = k]\,)\, \mathbb{P}_{[X=k]}(\, [T = j]\,)$.

Démonstration.
Par définition d'une probabilité conditionnelle :

$$\begin{aligned} & \sum_{j=1}^{n} \left(\sum_{k=1}^{n} j\ \mathbb{P}([X = k])\ \mathbb{P}_{[X=k]}([T = j]) \right) \\ = \ & \sum_{j=1}^{n} j \left(\sum_{k=1}^{n} \mathbb{P}([X = k] \cap [T = j]) \right) \\ = \ & \sum_{j=1}^{n} j\mathbb{P}([T = j]) \\ = \ & \mathbb{E}(T) \end{aligned}$$

(par formule des probabilités totales sur le système complet d'événements $\left(\ [X = k]\ \right)_{k\in[\![1,n]\!]}$)

$$\boxed{\mathbb{E}(T) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} j\ \mathbb{P}([X = k])\ \mathbb{P}_{[X=k]}([T = j]) \right)}$$

□

(iv) Démontrer : $\mathbb{E}(T) = \dfrac{\mathbb{E}(X + 1)}{2}$.

Démonstration.
D'après la question précédente :

$$\begin{aligned} \mathbb{E}(T) &= \sum_{j=1}^{n} \left(\sum_{k=1}^{n} j\ \mathbb{P}([X = k])\ \mathbb{P}_{[X=k]}([T = j]) \right) \\ &= \sum_{1\leqslant j,k\leqslant n} j\ \mathbb{P}([X = k])\ \mathbb{P}_{[X=k]}([T = j]) \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} j\ \mathbb{P}([X = k])\ \mathbb{P}_{[X=k]}([T = j]) \right) \\ &= \frac{\mathbb{E}(X + 1)}{2} \end{aligned}$$

*(d'après **4.a)(ii)**)*

$$\boxed{\mathbb{E}(T) = \frac{\mathbb{E}(X + 1)}{2}}$$

□

b) En réalité, en arrivant à l'arrêt de bus, on « tombe » dans un intervalle entre deux bus de manière proportionnelle à sa taille (plus l'intervalle est long, plus on a de chances de « tomber » dedans) : l'intervalle de temps est X^*, suivant la loi de X biaisée par la taille. Le temps d'attente T^* vérifie donc en fait, pour tout $k \in \{1, \ldots, n\}$:

$$\mathbb{P}_{[X^*=k]}(\, [T^* = j]\,) \ = \ \begin{cases} \dfrac{1}{k} & \text{si } j \in \{1, \ldots, k\} \\ 0 & \text{si } j > k \end{cases}$$

(i) Montrer que pour tout entier $k \in \{1, \ldots, n\}$, on a : $\sum\limits_{j=1}^{n} j\, \mathbb{P}_{[X^*=k]}(\, [T^* = j]\,) = \dfrac{k+1}{2}$.

Démonstration.
Soit $k \in [\![1, n]\!]$.
La loi de T^* conditionnellement à $[X^* = k]$ est exactement la même que la loi de T conditionnellement à $[X = k]$ (loi donnée en question ***4.a)***).

$$\boxed{\text{Avec les mêmes calculs qu'en } \textbf{\textit{4.a)(i)}} : \forall k \in [\![1, n]\!],\ \sum_{j=1}^{n} j\ \mathbb{P}_{[X=k]}([T = j]) = \frac{k+1}{2}.}$$

□

(ii) Démontrer : $\mathbb{E}(T^*) = \sum\limits_{j=1}^{n} \sum\limits_{k=1}^{n} j\, \mathbb{P}(\,[X^* = k]\,)\, \mathbb{P}_{[X^*=k]}(\,[T^* = j]\,)$.

Démonstration.
On effectue un raisonnement tout à fait similaire à celui de la question ***4.a)(iii)*** en appliquant la formule des probabilités totales sur le système complet d'événements $\left(\,[X^* = k]\,\right)_{k \in [\![1,n]\!]}$.

$$\boxed{\mathbb{E}(T^*) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} j\, \mathbb{P}(\,[X^* = k]\,)\, \mathbb{P}_{[X^*=k]}(\,[T^* = j]\,) \right)}$$

☐

(iii) Démontrer : $\mathbb{E}(T^*) = \dfrac{\mathbb{E}(X^* + 1)}{2}$.

Démonstration.
On effectue exactement le même raisonnement qu'en question ***4.a)(iv)***.

$$\boxed{\mathbb{E}(T^*) = \frac{\mathbb{E}(X^* + 1)}{2}}$$

☐

(iv) En déduire qu'on a : $\mathbb{E}(T^*) \geqslant \mathbb{E}(T)$.

Démonstration.

- Tout d'abord :

$$\begin{array}{rcll}
\mathbb{E}(T^*) \geqslant \mathbb{E}(T) & \Leftrightarrow & \dfrac{\mathbb{E}(X^*+1)}{2} \geqslant \dfrac{\mathbb{E}(X+1)}{2} & \textit{(d'après } \textbf{\textit{4.a)(iv)}} \textit{ et } \textbf{\textit{4.b)(iii)}}\textit{)} \\
& \Leftrightarrow & \mathbb{E}(X^*+1) \geqslant \mathbb{E}(X+1) & \\
& \Leftrightarrow & \mathbb{E}(X^*)+1 \geqslant \mathbb{E}(X)+1 & \textit{(par linéarité de l'espérance)} \\
& \Leftrightarrow & \mathbb{E}(X^*) \geqslant \mathbb{E}(X) &
\end{array}$$

- Démontrons alors : $\mathbb{E}(X^*) \geqslant \mathbb{E}(X)$.
  - × $X(\Omega) \subset [\![1,n]\!] \subset \mathbb{N}$
  - × Comme $X(\Omega) \subset [\![1,n]\!]$: $\mathbb{P}([X = 0]) = 0 < 1$.
  - × Comme la v.a.r. X est finie, elle admet un moment d'ordre 2 (et donc une espérance).

  On se place ainsi dans le cadre de la question ***2.***. On en déduit :

$$\mathbb{E}(X^*) \ \geqslant \ \mathbb{E}(X)$$

$$\boxed{\text{On en conclut : } \mathbb{E}(T^*) \geqslant \mathbb{E}(T).}$$

Commentaire

- Citer les hypothèses d'un théorème avant son utilisation est **indispensable**. Lorsque ce théorème n'est pas un résultat du cours mais un résultat démontré dans un sujet de concours, ce réflexe doit perdurer.
- On utilise par exemple dans cette question un résultat démontré dans une question antérieure (question ***2.***). On n'omettra donc en aucun cas de vérifier que l'on est placé dans le bon cadre d'application de cette propriété.

☐

Deuxième partie : Biais par la taille, propriétés

Dans cette partie, on démontre de nombreuses propriétés des variables aléatoires biaisées par la taille.

5. *Biais par la taille : le cas de variables à densité.*
Soit X une variable aléatoire **positive** de densité f et admettant une espérance $\mathbb{E}(X)$ strictement positive (donc $f(x) = 0$ pour tout x strictement négatif).
On définit la fonction g sur $\mathbb{R}$ par : $g : x \mapsto \dfrac{x}{\mathbb{E}(X)}\ f(x)$.

a) Montrer que g définit une densité d'une variable aléatoire positive.

Démonstration.

- La fonction g est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points car elle est le produit $g = g_1 \times f$ de :
  - × $g_1 : x \mapsto \dfrac{x}{\mathbb{E}(X)}$ qui est continue sur $\mathbb{R}$,
  - × f qui est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points.

  $\boxed{\text{La fonction } g \text{ est continue sur } \mathbb{R} \text{ sauf éventuellement en un nombre fini de points.}}$

- Soit $x \in \mathbb{R}$. Deux cas se présentent :
  - × si $x \in \,]-\infty, 0[$, alors :
  $$g(x) \ = \ \frac{x}{\mathbb{E}(X)}\ f(x) \ = \ \frac{x}{\mathbb{E}(X)} \times 0 \ = \ 0 \ \geqslant 0$$
  - × si $x \in [0, +\infty[$, alors, comme $x \geqslant 0$, $\mathbb{E}(X) > 0$ et $f(x) \geqslant 0$:
  $$g(x) \ = \ \frac{x}{\mathbb{E}(X)}\ f(x) \ \geqslant \ 0$$

  $\boxed{\text{On en déduit : } \forall x \in \mathbb{R},\ g(x) \geqslant 0.}$

- Démontrons que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} g(x)\ dx$ est convergente et vaut 1.
  On remarque, pour tout $x \in \mathbb{R}$:
  $$g(x) \ = \ \frac{x}{\mathbb{E}(X)}\ f(x) \ = \ \frac{1}{\mathbb{E}(X)}\ x\, f(x)$$
  Or la v.a.r. X admet une espérance. Ainsi l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x\, f(x)\ dx$ est convergente, et donc $\displaystyle\int_{-\infty}^{+\infty} g(x)\ dx$ aussi. *(on ne change pas la nature d'une intégrale en multipliant son intégrande par une constante non nulle)*
  De plus :
  $$\int_{-\infty}^{+\infty} g(x)\ dx \ = \ \int_{-\infty}^{+\infty} \frac{1}{\mathbb{E}(X)}\ x\, f(x)\ dx \ = \ \frac{1}{\mathbb{E}(X)} \int_{-\infty}^{+\infty} x\, f(x)\ dx \ = \ \frac{1}{\mathbb{E}(X)}\ \mathbb{E}(X) \ = \ 1$$

  $\boxed{\text{La fonction } g \text{ est donc une densité.}}$

$\boxed{\begin{array}{c}\text{Comme de plus } g \text{ est nulle en dehors de } [0, +\infty[\text{, on peut considérer que} \\ \text{c'est une densité d'une v.a.r. positive.}\end{array}}$ □

Soit une variable aléatoire X^* dont une densité est g. On dit que X^* suit *la loi de X biaisée par la taille.*

b) Soit a un réel strictement positif.

(i) Montrer que la variable aléatoire $a\,X$ possède pour densité $x \mapsto \dfrac{1}{a}\, f\left(\dfrac{x}{a}\right)$.

Démonstration.

- La v.a.r. $W = a\,X$ est une transformée affine de la v.a.r. à densité X.
On en conclut que la v.a.r. W est une v.a.r. à densité.
- Tout d'abord, comme $X(\Omega) \subset \mathbb{R}_+$ et $a > 0$, alors : $W(\Omega) = (a\,X)(\Omega) \subset \mathbb{R}_+$.

$$\boxed{W(\Omega) \subset \mathbb{R}_+}$$

- Soit $x \in \mathbb{R}$. Deux cas se présentent :
  × si $x \in \,]-\infty, 0[$, alors $[W \leqslant x] = \varnothing$ (car $W(\Omega) \subset \mathbb{R}_+$). Donc :

$$F_W(x) \;=\; \mathbb{P}([W \leqslant x]) \;=\; \mathbb{P}(\varnothing) \;=\; 0$$

  × si $x \in [0, +\infty[$, alors :

$$\begin{array}{rcll} F_W(x) & = & \mathbb{P}([a\,X \leqslant x]) & \\ & = & \mathbb{P}\left(\left[X \leqslant \dfrac{x}{a}\right]\right) & \textit{(car } a > 0\textit{)} \\ & = & F_X\left(\dfrac{x}{a}\right) & \end{array}$$

$$\boxed{\text{Finalement : } F_W : x \mapsto \left\{ \begin{array}{cl} 0 & \text{si } x < 0 \\ F_X\left(\dfrac{x}{a}\right) & \text{si } x \geqslant 0 \end{array} \right. .}$$

- La fonction F_W est continue sur $\mathbb{R}$ car elle est la composée $F_W = F_X \circ g_1$ où :
  × $g_1 : x \mapsto \dfrac{x}{a}$ est :
    - continue sur $\mathbb{R}$,
    - telle que : $g_1(\mathbb{R}) \subset \mathbb{R}$.

  × F_X est continue sur $\mathbb{R}$, car X est une v.a.r. à densité.
- La fonction F_W est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf (éventuellement) en un nombre fini de points avec des arguments similaires à ceux de la continuité sur $\mathbb{R}$.

$$\boxed{\text{On en déduit que } W \text{ est une v.a.r. à densité.}}$$

- Pour déterminer une densité f_W de Y, on dérive la fonction F_W en les points où elle est de classe $\mathcal{C}^1$. Soit $x \in \mathbb{R}$.
  × Si F_X est de classe $\mathcal{C}^1$ en x :

$$\begin{array}{rcl} f_W(x) & = & F_W'(x) \;=\; (F_X \circ g_1)'(x) \;=\; (F_X' \circ g_1)(x) \times g_1'(x) \\ & = & F_X'\big(g_1(x)\big) \times \dfrac{1}{a} \;=\; F_X'\left(\dfrac{x}{a}\right) \;=\; \dfrac{1}{a}\, f\left(\dfrac{x}{a}\right) \end{array}$$

  × Si F_X n'est pas de classe $\mathcal{C}^1$ en x, on choisit : $f_W(x) = \dfrac{1}{a}\, f\left(\dfrac{x}{a}\right)$.

$$\boxed{\text{Finalement : } f_W : x \mapsto \dfrac{1}{a}\, f\left(\dfrac{x}{a}\right).}$$

□

Commentaire

- Notons que l'expression de F_W sur $[0,+\infty[$ est aussi valide sur $]-\infty,0[$.

  En effet, pour tout $x \in\ ]-\infty,0[$, comme $a>0$, alors : $\dfrac{x}{a}<0$. Ainsi, comme $X(\Omega)\subset\mathbb{R}_+$:

$$F_X\left(\frac{x}{a}\right) \ = \ \mathbb{P}\left(\left[X\leqslant\frac{x}{a}\right]\right) \ = \ \mathbb{P}(\varnothing) \ = \ 0 \ = \ F_W(x)$$

  On opte cependant ici pour une présentation faisant intervenir une disjonction de cas pour mettre en valeur le fait que F_W s'annule en dehors de $[0,+\infty[$.
- On peut détailler le dernier point précédent concernant les points où F_X est de classe $\mathcal{C}^1$.
  - × Si F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, alors la fonction F_W est également de classe $\mathcal{C}^1$ sur $\mathbb{R}$ en tant composée $F_W=F_X\circ g_1$ (même démonstration que la continuité de F_W sur $\mathbb{R}$).
  - × Si F_X n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}$ tout en entier, on peut déterminer plus précisément les points où la fonction F_W n'est pas de classe $\mathcal{C}^1$.
    - Comme X est une v.a.r. à densité, sa fonction de répartition F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points. Notons $\{x_1,\dots,x_n\}$ l'ensemble des points où la fonction F_X n'est pas de classe $\mathcal{C}^1$.
    - Comme $F_W : x\mapsto F_X\left(\dfrac{x}{a}\right)$, on en déduit que la fonction F_W est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en $a\,x_1$, …, $a\,x_n$.

(ii) En déduire que $(a\,X)^*$ et $a\cdot X^*$ possèdent la même loi.

Démonstration.

- Tout d'abord, d'après la question précédente, la v.a.r. $W=a\,X$ est :
  - × positive. En effet : $W(\Omega)\subset\mathbb{R}_+$.
  - × de densité : $f_W : x\mapsto \dfrac{1}{a}\ f\left(\dfrac{x}{a}\right)$.
  - × admettant une espérance, en tant que transformée affine de X qui en admet une.
    De plus, par linéarité de l'espérance :

$$\mathbb{E}(W) \ = \ \mathbb{E}(a\,X) \ = \ a\,\mathbb{E}(X) \ > \ 0 \quad \textit{(car } a>0 \textit{ et } \mathbb{E}(X)>0\textit{)}$$

$$\boxed{\text{La v.a.r. } W^*=(a\,X)^* \text{ est donc bien définie.}}$$

- Par définition de W^*, cette v.a.r. admet pour densité la fonction $h_1 : x\mapsto \dfrac{x}{\mathbb{E}(W)}\ f_W(x)$.

  Or, pour tout $x\in\mathbb{R}$:

$$h_1(x) \ = \ \frac{x}{\mathbb{E}(W)}\ f_W(x) \ = \ \frac{x}{a\,\mathbb{E}(X)}\times\frac{1}{a}\ f\left(\frac{x}{a}\right) \ = \ \frac{x}{a^2\,\mathbb{E}(X)}\ f\left(\frac{x}{a}\right)$$

$$\boxed{\text{Une densité de } (a\,X)^* \text{ est } h_1 : x\mapsto \frac{x}{a^2\,\mathbb{E}(X)}\ f\left(\frac{x}{a}\right).}$$

- En procédant comme en question ***5.b)(i)***, on démontre que $a \cdot X^*$ est une v.a.r. à densité de densité $h_2 : x \mapsto \dfrac{1}{a}\, g\left(\dfrac{x}{a}\right)$. Ainsi, pour tout $x \in \mathbb{R}$:

$$h_2(x) \ = \ \frac{1}{a}\, g\left(\frac{x}{a}\right) \ = \ \frac{1}{a} \times \frac{\frac{x}{a}}{\mathbb{E}(X)}\, f\left(\frac{x}{a}\right) \ = \ \frac{x}{a^2\, \mathbb{E}(X)}\, f\left(\frac{x}{a}\right)$$

$$\boxed{\text{Une densité de } a \cdot X^* \text{ est } h_2 : x \mapsto \frac{x}{a^2\, \mathbb{E}(X)}\, f\left(\frac{x}{a}\right).}$$

Or une densité caractérise la loi d'une v.a.r. . On en déduit que les v.a.r. $(a\, X)^*$ et $a \cdot X^*$ suivent la même loi.

Commentaire

- Usuellement, on utilise la propriété au programme suivante : « la fonction de répartition caractérise la loi ». Dans cette question, on privilégie l'utilisation du corollaire : « une densité caractérise la loi ». Ou plus précisément, pour toute v.a.r. U et V à densité :

$$f_U = f_V \qquad \Rightarrow \qquad U \text{ et } V \text{ ont même loi}$$

- Démontrons cette proposition.
Supposons : $f_U = f_V$. Alors, pour tout $x \in \mathbb{R}$:

$$F_U(x) \ = \ \int_{-\infty}^{x} f_U(t)\, dt \ = \ \int_{-\infty}^{x} f_V(t)\, dt \ = \ F_V(x)$$

Les v.a.r. U et V ont donc même fonction de répartition. Or la fonction de répartition caractérise la loi. Ainsi U et V ont même loi.

□

c) *Une propriété importante.*
Soit $h : [0, +\infty[\to \mathbb{R}$ une fonction bornée et continue sauf éventuellement en un nombre fini de points. Montrer que $\mathbb{E}\big(X\, h(X)\big)$ est bien défini et :

$$\mathbb{E}\big(h(X^*)\big) \ = \ \frac{1}{\mathbb{E}(X)}\, \mathbb{E}\big(X\, h(X)\big)$$

Démonstration.

- Par théorème de transfert, la v.a.r. $X\, h(X)$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_0^{+\infty} x\, h(x)\, f(x)\, dx$ est absolument convergente.
- De plus, comme la fonction h est bornée, alors il existe $M \geqslant 0$ tel que :

$$\forall x \in [0, +\infty[, \quad |h(x)| \ \leqslant \ M$$

Soit $x \in [0, +\infty[$.

$$\begin{array}{rrcll} \text{comme} & |h(x)| & \leqslant & M & \\ \text{alors} & |h(x)|\, x\, f(x) & \leqslant & M\, x\, f(x) & \textit{(car : } x\, f(x) \geqslant 0\textit{)} \\ \text{donc} & |h(x)\, x\, f(x)| & \leqslant & M\, x\, f(x) & \textit{(toujours car : } x\, f(x) \geqslant 0\textit{)} \end{array}$$

Ainsi :

× $\forall x \in [0, +\infty[,\ 0 \leqslant |h(x)\, x\, f(x)| \leqslant M\, x\, f(x)$

× l'intégrale $\displaystyle\int_0^{+\infty} x\, f(x)\ dx$ est convergente car la v.a.r. X admet une espérance. Donc l'intégrale $\displaystyle\int_0^{+\infty} M\, x\, f(x)\ dx$ aussi.

Par critère de comparaison d'intégrales généralisées de fonctions continues positives, l'intégrale $\displaystyle\int_0^{+\infty} |h(x)\, x\, f(x)|\ dx$ est convergente. On en déduit que l'intégrale $\displaystyle\int_0^{+\infty} h(x)\, x\, f(x)\ dx$ est absolument convergente.

$$\boxed{\text{La v.a.r. } X\, h(X) \text{ admet donc une espérance.}}$$

- Par théorème de transfert, la v.a.r. $h(X^*)$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_0^{+\infty} h(x)\, g(x)\ dx$ est absolument convergente.
- Soit $x \in [0, +\infty[$.

$$|h(x)\, g(x)| \ = \ |h(x) \times \frac{x}{\mathbb{E}(X)}\ f(x)| \ = \ \frac{1}{\mathbb{E}(X)}\ |h(x)\, x\, f(x)| \qquad \textit{(car } \mathbb{E}(X) > 0\textit{)}$$

Or, d'après ce qui précède, l'intégrale $\displaystyle\int_0^{+\infty} h(x)\, x\, f(x)\ dx$ est absolument convergente. On en déduit que $\displaystyle\int_0^{+\infty} h(x)\, g(x)\ dx$ est absolument convergente.

$$\boxed{\text{La v.a.r. } h(X^*) \text{ admet donc une espérance.}}$$

- Enfin :

$$\begin{aligned}
\mathbb{E}\big(h(X^*)\big) &= \int_0^{+\infty} h(x)\, g(x)\ dx && \textit{(par théorème de transfert)} \\
&= \int_0^{+\infty} \frac{1}{\mathbb{E}(X)}\ h(x)\, x\, f(x)\ dx \\
&= \frac{1}{\mathbb{E}(X)} \int_0^{+\infty} h(x)\, x\, f(x)\ dx && \textit{(par linéarité de l'intégrale, car les intégrales en présence sont convergentes)} \\
&= \frac{1}{\mathbb{E}(X)}\ \mathbb{E}\big(X\, h(X)\big) && \textit{(par théorème de transfert)}
\end{aligned}$$

$$\boxed{\mathbb{E}\big(h(X^*)\big) = \frac{1}{\mathbb{E}(X)}\ \mathbb{E}\big(X\, h(X)\big)}$$

□

On pose alors la définition suivante (**que la variable X soit à densité ou non**) : si X est une variable aléatoire réelle positive d'espérance $\mathbb{E}(X)$ strictement positive, on dit que **la variable aléatoire positive Y suit la loi de X biaisée par la taille** si on a :

$$\mathbb{E}\big(h(Y)\big) \ = \ \frac{1}{\mathbb{E}(X)}\ \mathbb{E}\big(X\, h(X)\big)$$

pour toute fonction $h : [0, +\infty[\ \to \mathbb{R}$ bornée et continue sauf éventuellement en un nombre fini de points.

Commentaire

- L'énoncé fait ici une petite erreur de frappe en adoptant la notation Y pour la v.a.r. positive suivant la loi de X biaisée par la taille, au lieu de la notation X^*.
- On prendra garde aux sous-entendus fournis par cette définition. On comprend en particulier que, si X est une v.a.r. positive d'espérance $\mathbb{E}(X)$ strictement positive, alors, pour toute fonction $h : [0, +\infty[\to \mathbb{R}$ bornée et continue sauf éventuellement en un nombre fini de points, les v.a.r. $h(X^*)$ et $X\, h(X)$ admettent une espérance. On utilisera plusieurs fois ce résultat par la suite.

6. Dans cette question, on se fixe $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ deux fonctions **croissantes**. Soit X une variable aléatoire telle que les espérances $\mathbb{E}\big(f(X)\big)$, $\mathbb{E}\big(g(X)\big)$ et $\mathbb{E}\big(f(X)\, g(X)\big)$ sont bien définies.

a) Montrer que quels que soient les réels x_1 et x_2, on a : $\big(f(x_1) - f(x_2)\big)\, \big(g(x_1) - g(x_2)\big) \geqslant 0$.

Démonstration.
Soit $(x_1, x_2) \in \mathbb{R}^2$. Deux cas se présentent.

- si $x_1 \leqslant x_2$, alors, comme f et g sont croissantes sur $\mathbb{R}$:

$$f(x_1) \leqslant f(x_2) \qquad \text{et} \qquad g(x_1) \leqslant g(x_2)$$

  On en déduit :

$$f(x_1) - f(x_2) \leqslant 0 \qquad \text{et} \qquad g(x_1) - g(x_2) \leqslant 0$$

  Ainsi :

$$\big(f(x_1) - f(x_2)\big)\, \big(g(x_1) - g(x_2)\big) \geqslant 0$$

- si $x_1 > x_2$, alors, comme f et g sont croissantes sur $\mathbb{R}$:

$$f(x_1) \geqslant f(x_2) \qquad \text{et} \qquad g(x_1) \geqslant g(x_2)$$

  On en déduit :

$$f(x_1) - f(x_2) \geqslant 0 \qquad \text{et} \qquad g(x_1) - g(x_2) \geqslant 0$$

  Ainsi :

$$\big(f(x_1) - f(x_2)\big)\, \big(g(x_1) - g(x_2)\big) \geqslant 0$$

$$\boxed{\text{Finalement : } \forall (x_1, x_2) \in \mathbb{R}^2,\ \big(f(x_1) - f(x_2)\big)\, \big(g(x_1) - g(x_2)\big) \geqslant 0.}$$

□

b) Soient X_1 et X_2 deux variables aléatoires indépendantes, de même loi que X. Démontrer :

$$\mathbb{E}\Big(\big(f(X_1) - f(X_2)\big)\, \big(g(X_1) - g(X_2)\big)\Big) = 2\, \mathbb{E}\big(f(X)\, g(X)\big) - 2\, \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big)$$

Démonstration.

- Tout d'abord :

$$\big(f(X_1) - f(X_2)\big)\, \big(g(X_1) - g(X_2)\big) = f(X_1)\, g(X_1) - f(X_1)\, g(X_2) - f(X_2)\, g(X_1) + f(X_2)\, g(X_2)$$

- Démontrons que les 4 v.a.r. précédentes admettent une espérance.
  - × Comme les v.a.r. X_1 et X_2 sont indépendantes, alors, par lemme des coalitions, les v.a.r. $f(X_1)$ et $g(X_2)$ sont indépendantes. Ainsi :
    - les v.a.r. $f(X_1)$ et $g(X_2)$ sont indépendantes,
    - la v.a.r. $f(X_1)$ admet une espérance, car elle suit la même loi que $f(X)$ qui en admet une,
    - la v.a.r. $g(X_2)$ admet une espérance, car elle suit la même loi que $g(X)$ qui en admet une.

    On en déduit que la v.a.r. $f(X_1)\, g(X_2)$ admet une espérance.

  - × Comme les v.a.r. X_1 et X_2 sont indépendantes, alors, par lemme des coalitions, les v.a.r. $g(X_1)$ et $f(X_2)$ sont indépendantes. Ainsi :
    - les v.a.r. $g(X_1)$ et $f(X_2)$ sont indépendantes,
    - la v.a.r. $g(X_1)$ admet une espérance, car elle suit la même loi que $g(X)$ qui en admet une,
    - la v.a.r. $f(X_2)$ admet une espérance, car elle suit la même loi que $f(X)$ qui en admet une.

    On en déduit que la v.a.r. $g(X_1)\, f(X_2)$ admet une espérance.

  - × Les v.a.r. $f(X_1)\, g(X_1)$ et $f(X_2)\, g(X_2)$ ont même loi que la v.a.r. $f(X)\, g(X)$ admet une espérance.

    Les v.a.r. $f(X_1)\, g(X_1)$ et $f(X_2)\, g(X_2)$ admettent une espérance.

  Avec le 1er point, on en déduit que la v.a.r. $\big(f(X_1) - f(X_2)\big)\, \big(g(X_1) - g(X_2)\big)$ admet une espérance.

- De plus :

$$\begin{array}{rll} & \mathbb{E}\Big(\big(f(X_1) - f(X_2)\big)\, \big(g(X_1) - g(X_2)\big)\Big) & \\ = & \mathbb{E}\big(f(X_1)\, g(X_1) - f(X_1)\, g(X_2) - f(X_2)\, g(X_1) + f(X_2)\, g(X_2)\big) & \\ = & \mathbb{E}\big(f(X_1)\, g(X_1)\big) - \mathbb{E}\big(f(X_1)\, g(X_2)\big) - \mathbb{E}\big(f(X_2)\, g(X_1)\big) + \mathbb{E}\big(f(X_2)\, g(X_2)\big) & \textit{(par linéarité de l'espérance)} \\ = & \mathbb{E}\big(f(X_1)\, g(X_1)\big) - \mathbb{E}\big(f(X_1)\big)\, \mathbb{E}\big(g(X_2)\big) - \mathbb{E}\big(f(X_2)\big)\, \mathbb{E}\big(g(X_1)\big) + \mathbb{E}\big(f(X_2)\, g(X_2)\big) & \textit{(*)} \\ = & \mathbb{E}\big(f(X)\, g(X)\big) - \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big) - \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big) + \mathbb{E}\big(f(X)\, g(X)\big) & \textit{(car } X_1 \textit{ et } X_2 \textit{ ont même loi que } X\textit{)} \\ = & 2\, \mathbb{E}\big(f(X)\, g(X)\big) - 2\, \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big) & \end{array}$$

  L'égalité $(*)$ est obtenue par indépendance de $f(X_1)$ et $g(X_2)$, puis de $f(X_2)$ et $g(X_1)$.

$$\boxed{\mathbb{E}\Big(\big(f(X_1) - f(X_2)\big)\, \big(g(X_1) - g(X_2)\big)\Big) = 2\, \mathbb{E}\big(f(X)\, g(X)\big) - 2\, \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big)}$$

□

c) En déduire : $\mathbb{E}\big(f(X)\,g(X)\big) \geqslant \mathbb{E}\big(f(X)\big)\,\mathbb{E}\big(g(X)\big)$.

Démonstration.

- On sait que la v.a.r. $Z = \big(f(X_1) - f(X_2)\big)\,\big(g(X_1) - g(X_2)\big)$:
  - × est une v.a.r. positive, d'après ***6.a)***
  - × admet une espérance d'après la question précédente.

  Par positivité de l'espérance : $\mathbb{E}(Z) \geqslant 0$, *i.e.*
  $$\mathbb{E}\Big(\big(f(X_1) - f(X_2)\big)\,\big(g(X_1) - g(X_2)\big)\Big) \geqslant 0$$
- D'après la question précédente :
  $$2\,\mathbb{E}\big(f(X)\,g(X)\big) - 2\,\mathbb{E}\big(f(X)\big)\,\mathbb{E}\big(g(X)\big) \ \geqslant\ 0$$
  Ainsi :
  $$2\,\mathbb{E}\big(f(X)\,g(X)\big) \ \geqslant\ 2\,\mathbb{E}\big(f(X)\big)\,\mathbb{E}\big(g(X)\big)$$

$$\boxed{\text{Finalement : } \mathbb{E}\big(f(X)\,g(X)\big) \geqslant \mathbb{E}\big(f(X)\big)\,\mathbb{E}\big(g(X)\big).}$$

□

7. Dans cette question, on suppose que X est une variable aléatoire positive d'espérance strictement positive, et telle que $\mathbb{E}(X^{m+1})$ existe pour un entier $m \geqslant 1$ donné.

a) Soit p un entier naturel tel que $1 \leqslant p \leqslant m$.

(i) Montrer que pour tout réel $x \geqslant 0$, on a : $0 \leqslant x^p \leqslant 1 + x^{m+1}$.

Démonstration.
Soit $x \in [0, +\infty[$. Deux cas se présentent :

- si $x \in [0, 1]$, alors :
  $$\begin{array}{lccccccl} & 0 & \leqslant & x & \leqslant & 1 & & \\ \text{donc} & 0 & \leqslant & x^p & \leqslant & 1 & & \textit{(par croissance de } x \mapsto x^p \textit{ sur } \mathbb{R}_+\textit{)} \\ \text{d'où} & 0 & \leqslant & x^p & \leqslant & 1 & \leqslant\ 1 + x^{m+1} & \textit{(car } x^{m+1} \geqslant 0\textit{)} \end{array}$$
- si $x \in \,]1, +\infty[$, alors :
  $$\begin{array}{lcccl} & p & < & m+1 & \\ \text{donc} & p\,\ln(x) & < & (m+1)\,\ln(x) & \textit{(car, comme } x > 1 \textit{ : } \ln(x) > 0\textit{)} \\ \text{d'où} & x^p & \leqslant & x^{m+1} & \textit{(par croissance de } \exp \textit{ sur } \mathbb{R}\textit{)} \\ \text{ainsi} & x^p & \leqslant & x^{m+1} \ \leqslant\ 1 + x^{m+1} & \end{array}$$
  De plus, comme $x > 1$, alors, par croissance de $x \mapsto x^p$ sur $\mathbb{R}_+$: $x^p \geqslant 1 \geqslant 0$. Ainsi :
  $$0 \ \leqslant\ x^p \ \leqslant\ 1 + x^{m+1}$$

$$\boxed{\text{Finalement : } \forall x \in [0, +\infty[,\ 0 \leqslant x^p \leqslant 1 + x^{m+1}.}$$

□

(ii) Montrer que $\mathbb{E}(X^p)$ existe.

Démonstration.

Commentaire

Cette question nécessite l'utilisation d'un théorème dit de « domination ». Ce théorème s'énonce de la façon suivante :

Soient U et V deux variables aléatoires.

Supposons :

× $0 \leqslant |U| \leqslant V$,

× la v.a.r. V admet une espérance.

Alors la v.a.r. U admet une espérance et : $|\mathbb{E}(U)| \leqslant \mathbb{E}(V)$.
Ce théorème ne fait pas partie du programme ECE (il fait partie du programme ECS).
Nous utiliserons dans ce sujet ce résultat sans démonstration.

- Comme X est une v.a.r. positive, d'après ***7.a)(i)*** :

$$0 \ \leqslant \ X^p \ \leqslant \ 1 + X^{m+1}$$

- De plus, la v.a.r. $1 + X^{m+1}$ admet une espérance en tant que transformée affine de X^{m+1} qui en admet une.

$$\boxed{\text{Par existence de l'espérance par domination, la v.a.r. } X^p \text{ admet une espérance.}}$$

Commentaire

En considérant que la v.a.r. X était soit discrète, soit à densité, il était possible de résoudre cette question sans utiliser de résultat hors programme. Détaillons cela.

- Cas d'une v.a.r. discrète : supposons $X(\Omega) = \{x_k \mid k \in \mathbb{N}\} \subset \mathbb{R}_+$.

× La v.a.r. X admet un moment d'ordre p si et seulement si la série $\sum\limits_{k \geqslant 0} x_k^p \, \mathbb{P}([X = x_k])$ est absolument convergente. Comme X est une v.a.r. positive, cette série est à termes positifs. Cela revient donc à démontrer que cette série est convergente.

× D'après la question précédente, pour tout $k \in \mathbb{N}$, comme $x_k \geqslant 0$ (car X est positive) :

$$\begin{array}{lccccl} & 0 & \leqslant & x_k^p & \leqslant & 1 + x_k^{m+1} \\ \text{donc} & 0 & \leqslant & x_k^p \, \mathbb{P}([X = x_k]) & \leqslant & (1 + x_k^{m+1}) \, \mathbb{P}([X = x_k]) \qquad \textit{(car } \mathbb{P}([X = x_k]) \geqslant 0\textit{)} \\ \text{d'où} & 0 & \leqslant & x_k^p \, \mathbb{P}([X = x_k]) & \leqslant & \mathbb{P}([X = x_k]) + x_k^{m+1} \, \mathbb{P}([X = x_k]) \end{array}$$

× On obtient ainsi :

- $\forall k \in \mathbb{N}$, $0 \leqslant x_k^p \, \mathbb{P}([X = x_k]) \leqslant \mathbb{P}([X = x_k]) + x_k^{m+1} \, \mathbb{P}([X = x_k])$
- la série $\sum\limits_{k \geqslant 0} \mathbb{P}([X = x_k])$ est convergente car $\big([X = x_k] \big)_{k \in \mathbb{N}}$ forme un système complet d'événements, et la série $\sum\limits_{k \geqslant 0} x_k^{m+1} \, \mathbb{P}([X = x_k])$ est convergente car X admet un moment d'ordre $m+1$ (d'après l'énoncé).
Ainsi la série $\sum\limits_{k \geqslant 0} \Big(\mathbb{P}([X = x_k]) + x_k^{m+1} \, \mathbb{P}([X = x_k]) \Big)$ est convergente.

Par critère de comparaison de séries à termes positifs, la série $\sum\limits_{k \geqslant 0} x_k^p \, \mathbb{P}([X = x_k])$ est convergente. La v.a.r. X^p admet donc une espérance.

Commentaire

- Cas d'une v.a.r. à densité : supposons que X est une v.a.r. à densité.
  Comme X est une v.a.r. positive, elle admet une densité f nulle en dehors de $[0, +\infty[$.
  × La v.a.r. X admet un moment d'ordre p si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} t^p f(t)\, dt$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour un calcul de moments du type $\displaystyle\int_{-\infty}^{+\infty} t^r f(t)\, dt$.
  × Comme f est nulle en dehors de $[0, +\infty[$:
  $$\int_{-\infty}^{+\infty} t^p f(t)\, dt = \int_0^{+\infty} t^p f(t)\, dt$$
  × D'après la question précédente, pour tout $t \in [0, +\infty[$:
  $$\begin{array}{rcccccl} & 0 & \leqslant & t^p & \leqslant & 1 + t^{m+1} & \\ \text{donc} & 0 & \leqslant & t^p f(t) & \leqslant & (1 + t^{m+1}) f(t) & \textit{(car } f(t) \geqslant 0\textit{)} \\ \text{d'où} & 0 & \leqslant & t^p f(t) & \leqslant & f(t) + t^{m+1} f(t) & \end{array}$$
  × On obtient ainsi :
  - $\forall x \in [0, +\infty[,\ 0 \leqslant t^p f(t) \leqslant f(t) + t^{m+1} f(t)$
  - l'intégrale $\displaystyle\int_0^{+\infty} f(t)\, dt$ est convergente car c'est le moment d'ordre 0 de la v.a.r. X et l'intégrale $\displaystyle\int_0^{+\infty} t^{m+1} f(t)\, dt$ est convergente car c'est le moment d'ordre $m+1$ de X (qui existe d'après l'énoncé). Ainsi l'intégrale $\displaystyle\int_0^{+\infty} f(t) + t^{m+1} f(t)\, dt$ est convergente.

  Par critère de comparaison d'intégrales généralisées de fonctions continues positives, l'intégrale $\displaystyle\int_0^{+\infty} t^p f(t)\, dt$ est convergente. La v.a.r. X^p admet donc une espérance.

□

b) Démontrer : $\mathbb{E}(X^{m+1}) \geqslant \mathbb{E}(X)\,\mathbb{E}(X^m)$.

Démonstration.
On souhaite appliquer la question ***6.***

- On considère les fonctions :

$$f : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x \end{array} \qquad \text{et} \qquad g : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \left\{\begin{array}{ll} 0 & \text{si } x < 0 \\ x^m & \text{si } x \geqslant 0 \end{array}\right. \end{array}$$

Commentaire

- Une première idée est d'appliquer la question ***6.*** à la fonction $g : x \mapsto x^m$. Malheureusement, c'est fonction n'est pas croissante sur $\mathbb{R}$ en toute généralité (par exemple, la fonction $x \mapsto x^2$ n'est pas croissante sur $]-\infty, 0[$).
- Pour pallier ce problème, on utilise la fait que la v.a.r. X est positive. Ainsi, la manière de définir la fonction g sur $]-\infty, 0[$ n'a pas grande importance. On choisit donc une fonction g :
  × qui coïncide avec la fonction $x \mapsto x^m$ sur $\mathbb{R}_+$,
  × dont l'expression est très simple sur $]-\infty, 0[$,
  × qui est croissante sur $\mathbb{R}$.

  C'est le cas de la fonction g proposée ici mais il y a avait bien d'autres choix possibles.

- On sait alors que :
  × les fonctions f et g sont croissantes sur $\mathbb{R}$,
  × la v.a.r. $f(X) = X$ admet une espérance,
  × la v.a.r. $g(X) = X^m$ admet une espérance d'après ***7.a)*** (car $m \in [\![1, m]\!]$).
  Démontrons qu'on a effectivement : $g(X) = X^m$.
  Soit $\omega \in \Omega$. Comme X est une v.a.r. positive, alors : $X(\omega) \geqslant 0$. Ainsi, par définition de g :

$$\big(g(X)\big)(\omega) \ = \ g\big(X(\omega)\big) \ = \ \big(X(\omega)\big)^m \ = \ X^m(\omega)$$

  × la v.a.r. $f(X)\, g(X) = X \times X^m = X^{m+1}$ admet une espérance, d'après l'énoncé.

  On est donc bien dans le cadre d'application de la question ***6.*** On en conclut :

$$\begin{array}{ccc} \mathbb{E}\big(f(X)\, g(x)\big) & \geqslant & \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big) \\ \shortparallel & & \shortparallel \\ \mathbb{E}(X^{m+1}) & & \mathbb{E}(X)\, \mathbb{E}(X^m) \end{array}$$

$$\boxed{\mathbb{E}(X^{m+1}) \geqslant \mathbb{E}(X)\, \mathbb{E}(X^m)}$$

Commentaire

- Citer les hypothèses d'un théorème avant son utilisation est **indispensable**. Lorsque ce théorème n'est pas un résultat du cours mais un résultat démontré dans un sujet de concours, ce réflexe doit perdurer.
- On utilise par exemple dans cette question un résultat démontré dans une question antérieure (question ***6.***). On n'omettra donc en aucun cas de vérifier que l'on est placé dans le bon cadre d'application de cette propriété.

□

c) En déduire : $\mathbb{E}\left((X^*)^m\right) \geqslant \mathbb{E}(X^m)$.

Commentaire

- Pour résoudre cette question, les concepteurs souhaitaient sans doute utiliser la définition de la v.a.r. X^* qui suit la loi de X biaisée par la taille. Autrement dit, pour toute fonction $h : [0, +\infty[\to \mathbb{R}$ bornée et continue sauf éventuellement en un nombre fini de points :

$$\mathbb{E}\big(h(X^*)\big) \;=\; \frac{1}{\mathbb{E}(X)}\ \mathbb{E}\big(X\, h(X)\big)$$

  L'idée étant d'appliquer cette égalité avec la fonction $h : x \mapsto x^m$. Cependant, cette fonction h n'est pas bornée sur $\mathbb{R}_+$ et cette définition ne peut donc être exploitée.

- Pour pouvoir démontrer le résultat demandé, on considèrera que la v.a.r. X est une v.a.r. à densité. Ne pas faire cette hypothèse rendait la résolution de cette question impossible dans le cadre du programme. Cette démarche n'est en aucun cas à généraliser. Les sujets sont d'habitude relus avec la plus grande attention et ne nécessitent pas l'introduction d'hypothèses personnelles.

Démonstration.
On considère que X est une v.a.r. de densité f.

- D'après la question ***5.a)***, la v.a.r. X^* est bien définie et est une v.a.r. de densité :

$$g : x \mapsto \frac{x}{\mathbb{E}(X)}\ f(x)$$

- La v.a.r. X^* admet donc un moment d'ordre m si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^m\, g(x)\ dx$ est absolument convergente, ce qui équivaut à démontrer sa convergence pour un calcul de moments du type $\displaystyle\int_{-\infty}^{+\infty} x^r\, g(x)\ dx$.

- On remarque, pour tout $x \in \mathbb{R}$:

$$x^m\, g(x) \;=\; x^m \times \frac{x}{\mathbb{E}(X)}\ f(x) \;=\; \frac{1}{\mathbb{E}(X)}\ x^{m+1}\, f(x)$$

  Or l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^{m+1}\, f(x)\ dx$ est convergente car la v.a.r. X admet un moment d'ordre $m+1$ d'après l'énoncé. On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^m\, g(x)\ dx$ est convergente.

$$\boxed{\text{La v.a.r. } X^* \text{ admet donc un moment d'ordre } m.}$$

- De plus :

$$\mathbb{E}\big((X^*)^m\big) \;=\; \int_{-\infty}^{+\infty} x^m\, g(x)\ dx \;=\; \frac{1}{\mathbb{E}(X)} \int_{-\infty}^{+\infty} x^{m+1}\, f(x)\ dx \;=\; \frac{1}{\mathbb{E}(X)}\ \mathbb{E}(X^{m+1})$$

$$\boxed{\mathbb{E}\big((X^*)^m\big) \;=\; \frac{1}{\mathbb{E}(X)}\ \mathbb{E}(X^{m+1})}$$

- Enfin, d'après la question précédente :

$$\begin{array}{lccl}
 & \mathbb{E}(X^{m+1}) & \geqslant & \mathbb{E}(X)\,\mathbb{E}(X^m) \\[2mm]
\text{donc} & \dfrac{1}{\mathbb{E}(X)}\ \mathbb{E}(X^{m+1}) & \geqslant & \mathbb{E}(X^m) \qquad \textit{(car } \mathbb{E}(X) > 0\textit{)} \\[2mm]
\text{d'où} & \mathbb{E}\big((X^*)^m\big) & \geqslant & \mathbb{E}(X^m) \qquad \textit{(d'après le point précédent)}
\end{array}$$

$$\boxed{\mathbb{E}\big((X^*)^m\big) \geqslant \mathbb{E}(X^m)}$$

Commentaire

- Avec des arguments similaires à ceux présentés dans la démonstration ci-dessus, on peut traiter cette question dans le cas où la v.a.r. X est discrète.
- En fait, il existe une théorie (largement hors programme) permettant d'unifier sous une même écriture l'espérance dans le cas discret et le cas continu. C'est sans doute ce que les concepteurs de ce sujet avaient en tête en demandant aux candidats de manipuler des v.a.r. quelconques.
- Sur les épreuves ESSEC II, il faut garder en tête que le barème est à la faveur du candidat : quand celui-ci montre qu'il sait répondre en partie à une question, il se verra allouer une bonne partie des points (voire tous les points dans certains cas).

$\square$

8. Pour A un événement, on note $\mathbb{1}_A$ la variable aléatoire définie par :

$$\mathbb{1}_A : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}$$

Pour tout t réel, on définit la fonction g_t sur $\mathbb{R}$ par : $g_t : x \mapsto \mathbb{1}_{]t,+\infty[}(x)$.

Commentaire

- L'énoncé introduit ici la définition de variable aléatoire indicatrice. Ce type de v.a.r. ne font pas partie du programme d'ECE. Donnons néanmoins certaines de leurs propriétés.
  - × Loi de $\mathbb{1}_A$.
    - Par définition de $\mathbb{1}_A$, cette v.a.r. ne prend comme valeur que 0 et 1.
      Donc $\mathbb{1}_A(\Omega) = \{0, 1\}$.
    - Soit $\omega \in \Omega$.
      $$\omega \in [\mathbb{1}_A = 1] \;\Leftrightarrow\; \mathbb{1}_A(\omega) = 1 \;\Leftrightarrow\; \omega \in A$$
      D'où : $[\mathbb{1}_A = 1] = A$. Ainsi : $\mathbb{P}([\mathbb{1}_A = 1]) = \mathbb{P}(A)$.

    On en déduit : $\mathbb{1}_A \hookrightarrow \mathcal{B}\left(\mathbb{P}(A)\right)$.
  - × En particulier :
    $$\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A) \qquad \text{et} \qquad \mathbb{V}(\mathbb{1}_A) = \mathbb{P}(A)\big(1 - \mathbb{P}(A)\big)$$
- Il peut aussi être utilise de savoir démontrer les propriétés suivantes.
  Soient B et C deux événements.
  - × $\mathbb{1}_{B\cap C} = \mathbb{1}_B \times \mathbb{1}_C$
  - × $\mathbb{1}_B + \mathbb{1}_{\overline{B}} = 1$

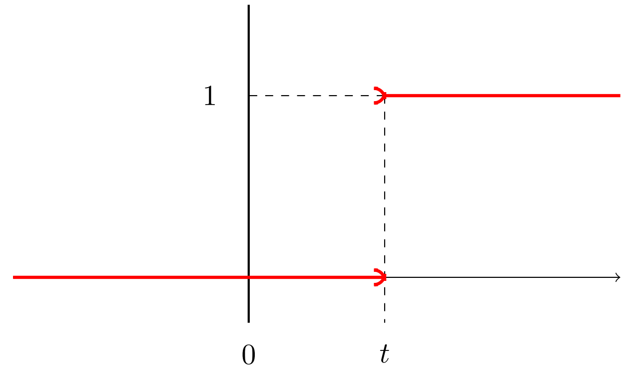
  Pour la démonstration de ces deux propriétés, on pourra, par exemple, se référer au sujet `ESSEC-II 2018`.

Commentaire

- La fonction g_t est quant à elle une fonction indicatrice. L'énoncé ne définit pas cette notation. Explicitons la :

$$g_t : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \left\{ \begin{array}{ll} 1 & \text{si } x \in \,]t, +\infty[\\ 0 & \text{sinon} \end{array} \right. \end{array}$$

Sa courbe représentative est donc assez simple.



- On prendra garde à ne pas confondre **variable aléatoire** indicatrice et **fonction** indicatrice. En effet :
  - × une variable aléatoire indicatrice est, bien évidemment, une variable aléatoire. En cette qualité, c'est une fonction de Ω dans $\mathbb{R}$.
  - × une fonction indicatrice est une fonction de $\mathbb{R}$ dans $\mathbb{R}$.

a) Montrer que la fonction $x \mapsto g_t(x)$ est croissante sur $\mathbb{R}$.

Démonstration.
Soit $(x, y) \in \mathbb{R}^2$. Supposons : $x \leqslant y$. Quatre cas se présentent.

- si $x \leqslant t$ et $y \leqslant t$, alors :
$$g_t(x) = 0 \qquad \text{et} \qquad g_t(y) = 0$$
D'où : $g_t(x) \leqslant g_t(y)$.
- si $x \leqslant t$ et $y > t$, alors :
$$g_t(x) = 0 \qquad \text{et} \qquad g_t(y) = 1$$
D'où : $g_t(x) \leqslant g_t(y)$.
- si $x > t$ et $y \leqslant t$, alors ce cas est impossible car on a supposé : $x \leqslant y$.
- si $x > t$ et $y > t$, alors :
$$g_t(x) = 1 \qquad \text{et} \qquad g_t(y) = 1$$
D'où : $g_t(x) \leqslant g_t(y)$.

Finalement, pour tout $(x, y) \in \mathbb{R}^2$: $x \leqslant y \quad \Rightarrow g_t(x) \leqslant g_t(y)$.

$$\boxed{\text{On en déduit que la fonction } g_t \text{ est croissante sur } \mathbb{R}.}$$

□

b) Soit X une variable aléatoire positive, admettant une espérance. Montrer que pour tout t réel, $\mathbb{E}\big(X\, g_t(X)\big)$ est bien défini et : $\mathbb{E}\big(X\, g_t(X)\big) \geqslant \mathbb{E}(X)\, \mathbb{P}(\,[X > t]\,)$.

Démonstration.

- On considère : $\mathbb{E}(X) > 0$.
  - × Alors la v.a.r. X^* est bien définie. On en déduit en particulier que, pour toute fonction h : $[0, +\infty[\ \to \mathbb{R}$ bornée et continue sauf éventuellement en un nombre fini de points, la v.a.r. $X\, h(X)$ admet une espérance.
  - × Or la fonction g_t est :
    - bornée. En effet, par définition, la fonction g_t est minorée par 0 et majorée par 1.
    - continue sur $\mathbb{R}$ sauf en t.

$$\boxed{\text{On en conclut que la v.a.r. } X\, g_t(X) \text{ admet une espérance.}}$$

Commentaire

- Dans cette question, on s'est permis d'écrire :

  « On considère : $\mathbb{E}(X) > 0$ »

  L'oubli de cette hypothèse n'était sans doute pas intentionnelle. Elle est en effet indispensable pour que la v.a.r. X^* de la question suivante soit bien définie.
- On pouvait répondre à cette question sans cette hypothèse de stricte positivité de l'espérance, mais au prix de l'utilisation du théorème hors programme d'existence de l'espérance par domination. Détaillons cette démonstration.
  - × Tout d'abord, on remarque :
  $$g_t(X) \ = \ \mathbb{1}_{[X>t]}$$
  En effet, pour tout $\omega \in \Omega$:
  $$\big(g_t(X)\big)(\omega) \ = \ g_t\big(X(\omega)\big) \ = \ \mathbb{1}_{]t,+\infty[}\big(X(\omega)\big)$$
  Deux cas se présentent :
    - si $X(\omega) \leqslant t$, alors :
    $$\mathbb{1}_{]t,+\infty[}\big(X(\omega)\big) \ = \ 0$$
    De plus, comme $X(\omega) \leqslant t$, alors : $\omega \in [X \leqslant t]$. D'où :
    $$\mathbb{1}_{[X>t]}(\omega) \ = \ 0$$
    - si $X(\omega) > t$, alors :
    $$\mathbb{1}_{]t,+\infty[}\big(X(\omega)\big) \ = \ 1$$
    De plus, comme $X(\omega) > t$, alors : $\omega \in [X > t]$. D'où :
    $$\mathbb{1}_{[X>t]}(\omega) \ = \ 1$$
  Finalement : $\forall \omega \in \Omega$, $\big(g_t(X)\big)(\omega) = \mathbb{1}_{[X>t]}(\omega)$.

Commentaire

× Par définition de $\mathbb{1}_{[X>t]}$, on obtient : $0 \leqslant g_t(X) \leqslant 1$. D'où, comme X est une v.a.r. positive :

$$0 \ \leqslant \ X\, g_t(X) \ \leqslant \ X$$

De plus, la v.a.r. X admet une espérance.
Par théorème d'existence de l'espérance par domination, la v.a.r. $X\, g_t(X)$ admet une espérance.

- On cherche à nouveau à appliquer la question ***6.*** On considère les fonctions :

$$f : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x \end{array} \qquad \text{et} \qquad g = g_t$$

  × Les fonctions f et g_t sont croissantes sur $\mathbb{R}$ (d'après ***8.a)*** pour la fonction g).
  × La v.a.r. $f(X) = X$ admet une espérance.
  × La v.a.r. $g_t(X)$ admet une espérance car c'est une v.a.r. finie. En effet :

$$\big(g_t(X)\big)(\Omega) = \mathbb{1}_{[X>t]}(\Omega) \subset \{0, 1\}$$

  × La v.a.r. $X\, g_t(X)$ admet une espérance d'après le point précédent.

  On est donc bien dans le cadre d'application de la question ***6.*** On en conclut :

$$\begin{array}{ccc} \mathbb{E}\big(f(X)\, g(X)\big) & \geqslant & \mathbb{E}\big(f(X)\big)\, \mathbb{E}\big(g(X)\big) \\ \shortparallel & & \shortparallel \\ \mathbb{E}\big(X\, g_t(X)\big) & & \mathbb{E}(X)\, \mathbb{E}\big(g_t(X)\big) \end{array}$$

- Enfin :
  × $\big(g_t(X)\big)(\Omega) \subset \{0, 1\}$.
  × de plus, par définition de $\mathbb{1}_{[X>t]}$:

$$\mathbb{P}([g_t(X) = 1]) \ = \ \mathbb{P}\big(\big[\mathbb{1}_{[X>t]} = 1\big]\big) \ = \ \mathbb{P}([X > t])$$

$$\boxed{\text{On en déduit : } g_t(X) \hookrightarrow \mathcal{B}\left(\mathbb{P}([X > t])\right).}$$

- En particulier : $\mathbb{E}\big(g_t(X)\big) = \mathbb{P}([X > t])$.

$$\boxed{\text{On obtient finalement : } \mathbb{E}\big(X\, g_t(X)\big) \geqslant \mathbb{E}(X)\, \mathbb{P}([X > t]).}$$

□

c) Démontrer, pour tout t réel : $\mathbb{P}(\,[X^* > t]\,) \geqslant \mathbb{P}(\,[X > t]\,)$.

On dit que X^* *domine stochastiquement* X.

Démonstration.
On considère toujours : $\mathbb{E}(X) > 0$.
(Rappelons que sans cette hypothèse, la v.a.r. X^* n'est pas définie)
Soit $t \in \mathbb{R}$. On note h_t la restriction de la fonction g_t à $[0, +\infty[$.

- Comme la fonction g_t est bornée et continue sur $\mathbb{R}$ sauf en t, alors la fonction $h_t : [0, +\infty[\to \mathbb{R}$ est bornée et continue sur $\mathbb{R}$ sauf éventuellement en t. Ainsi, par définition de X^* :

$$\mathbb{E}\big(g_t(X^*)\big) \ = \ \frac{1}{\mathbb{E}(X)}\, \mathbb{E}\big(X\, g_t(X)\big)$$

- Or, d'après la question ***8.b)*** :

$$\begin{array}{lrcll} & \mathbb{E}\big(X\, g_t(X)\big) & \geqslant & \mathbb{E}(X)\, \mathbb{P}([X > t]) & \\ \text{donc} & \dfrac{1}{\mathbb{E}(X)}\, \mathbb{E}\big(X\, g_t(X)\big) & \geqslant & \mathbb{P}([X > t]) & \textit{(car } \mathbb{E}(X) > 0\textit{)} \\ \text{d'où} & \mathbb{E}\big(g_t(X^*)\big) & \geqslant & \mathbb{P}([X > t]) & \end{array}$$

- Enfin, avec un raisonnement similaire à celui de la question précédente : $g_t(X^*) \hookrightarrow \mathcal{B}\left(\mathbb{P}([X^* > t])\right)$. D'où :

$$\mathbb{E}\big(g_t(X^*)\big) = \mathbb{P}([X^* > t])$$

$$\boxed{\text{Finalement : } \forall t \in \mathbb{R},\ \mathbb{P}([X^* > t]) \geqslant \mathbb{P}([X > t]).}$$ □

9. Soit $X_1, \ldots, X_n$ des variables aléatoires positives, indépendantes, non nécessairement de même loi. On suppose qu'elles admettent toutes une espérance strictement positive, et on note $\mu_i = \mathbb{E}(X_i)$. De plus, on pose : $\mu = \sum\limits_{i=1}^{n} \mu_i$ et $S_n = \sum\limits_{i=1}^{n} X_i$.

a) Donner $\mathbb{E}(S_n)$.

Démonstration.

- La v.a.r. S_n admet une espérance en tant que somme de v.a.r. qui en admettent une.
- De plus, par linéarité de l'espérance :

$$\mathbb{E}(S_n) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \sum_{i=1}^{n} \mu_i = \mu$$

$$\boxed{\mathbb{E}(S_n) = \mu}$$ □

b) Soit J une variable aléatoire à valeurs dans $\{1, \ldots, n\}$, de loi $\mathbb{P}(\,[J = k]\,) = \dfrac{\mu_k}{\mu}$. Quelle est la loi de J si les variables aléatoires X_i sont de même loi ?

Commentaire

On imagine que l'énoncé voulait donner la définition suivante pour la loi de la v.a.r. J :
- $J(\Omega) \subset [\![1, n]\!]$,
- $\forall k \in [\![1, n]\!]$, $\mathbb{P}([J = k]) = \dfrac{\mu_k}{\mu}$

Démonstration.

- D'après l'énoncé : $J(\Omega) \subset [\![1, n]\!]$.
- Soit $k \in [\![1, n]\!]$.

$$\mathbb{P}([J = k]) = \frac{\mu_k}{\mu}$$

Or :

× d'une part, comme $X_1, \ldots, X_n$ ont même loi :

$$\mu_k = \mathbb{E}(X_k) = \mathbb{E}(X_1) = \mu_1$$

× d'autre part, pour la même raison :

$$\mu \;=\; \sum_{i=1}^{n} \mathbb{E}(X_i) \;=\; \sum_{k=1}^{n} \mathbb{E}(X_1) \;=\; n\,\mathbb{E}(X_1) \;=\; n\,\mu_1$$

Ainsi : $\mathbb{P}([J = k]) \;=\; \dfrac{\mu_k}{\mu} \;=\; \dfrac{\cancel{\mu_1}}{n\,\cancel{\mu_1}} \;=\; \dfrac{1}{n}$.

$$\boxed{\text{On en déduit : } J \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket).}$$

□

On considère $X_1^*, \ldots, X_n^*$ des variables aléatoires indépendantes, indépendantes de $X_1, \ldots, X_n$ telles que, pour tout entier i tel que $1 \leqslant i \leqslant n$, X_i^* suive la loi de X_i biaisée par la taille.
Soit aussi J une variable aléatoire de loi $\mathbb{P}(\,[J = k]\,) = \dfrac{\mu_k}{\mu}$, indépendante de X_1, X_1^*, …, X_n, X_n^*.

On considère la variable aléatoire $X_J = \sum\limits_{j=1}^{n} X_j\, \mathbb{1}_{[J=j]}$ et on définit $T_n = S_n - X_J + X_J^*$. Autrement dit, on choisit un indice aléatoire J et, dans la somme $\sum\limits_{i=1}^{n} X_i$, on remplace X_J par X_J^*.

Commentaire

- En parlant d'« indice aléatoire » pour la variable aléatoire J, l'énoncé fait ici la confusion entre la variable aléatoire J et la valeur prise par celle-ci. Ainsi, dans l'écriture « X_J », J n'est en aucun cas un réel mais chaque chaque valeur prise par J en est un.
- La définition de la v.a.r. X_J est donc la suivante :

  si la v.a.r. J prend la valeur i, alors X_J prend la valeur X_i.

  Il faut ensuite comprendre que ce que l'énoncé définit comme X_J^* est :

  si la v.a.r. J prend la valeur i, alors X_J^* prend la valeur X_i^*.

  On aurait aimé que l'énoncé introduise cette v.a.r. X_J^* correctement de la façon suivante :

$$\begin{array}{rccc} X_J^* & : & \Omega & \to & \mathbb{R} \\ & & \omega & \mapsto & X_{J(\omega)}^*(\omega) \end{array}$$

- En exploitant la notion de v.a.r. indicatrice, on peut aussi définir la v.a.r. X_J^* ainsi :

$$X_J^* = \sum_{j=1}^{n} X_j^*\, \mathbb{1}_{[J=j]}$$

  On utilisera régulièrement cette définition par la suite.

c) Soit $h : [0, +\infty[\, \to \mathbb{R}$ une fonction bornée et continue sauf éventuellement en un nombre fini de points.

(i) Démontrer : $h(T_n) = \sum\limits_{i=1}^{n} h(T_n)\, \mathbb{1}_{[J=i]} = \sum\limits_{i=1}^{n} h(S_n - X_i - X_i^*)\, \mathbb{1}_{[J=i]}$.

Commentaire

L'énoncé de cette question comporte une erreur. En effet, il s'agit bien ici de démontrer :

$$h(T_n) = \sum_{i=1}^{n} h(T_n)\, \mathbb{1}_{[J=i]} = \sum_{i=1}^{n} h(S_n - X_i + X_i^*)\, \mathbb{1}_{[J=i]}$$

Démonstration.

- Tout d'abord :
$$\sum_{i=1}^{n} h(T_n)\, \mathbb{1}_{[J=i]} \ = \ h(T_n) \sum_{i=1}^{n} \mathbb{1}_{[J=i]}$$

- Démontrons alors : $\sum\limits_{i=1}^{n} \mathbb{1}_{[J=i]} = 1$.

  Soit $\omega \in \Omega$.

  × Comme $\big([J=i] \big)_{i\in[\![1,n]\!]}$ forme un système complet d'événements, alors : $\Omega = \bigcup\limits_{i=1}^{n} [J=i]$.

    Ainsi : $\omega \in \bigcup\limits_{i=1}^{n} [J=i]$. Il existe donc $i_0 \in [\![1,n]\!]$ tel que : $\omega \in [J=i_0]$.

  × On en déduit :

    - d'une part, par définition de $\mathbb{1}_{[J=i_0]}$:
    $$\mathbb{1}_{[J=i_0]}(\omega) \ = \ 1$$

    - d'autre part, pour tout $i \in [\![1,n]\!]$ tel que $i \neq i_0$, les événements $[J=i]$ et $[J=i_0]$ sont incompatibles. Ainsi : $\omega \notin [J=i]$. D'où :
    $$\mathbb{1}_{[J=i]}(\omega) \ = \ 0$$

  On obtient :
$$\sum_{i=1}^{n} \mathbb{1}_{[J=i]}(\omega) \ = \ \sum_{\substack{i=1 \\ i\neq i_0}}^{n} \mathbb{1}_{[J=i]}(\omega) + \mathbb{1}_{[J=i_0]}(\omega) \ = \ \cancel{\sum_{\substack{i=1 \\ i\neq i_0}}^{n} 0} \ + \ 1 \ = \ 1$$

$$\boxed{\text{On en déduit : } \sum_{i=1}^{n} h(T_n)\, \mathbb{1}_{[J=i]} \ = \ h(T_n) \sum_{i=1}^{n} \mathbb{1}_{[J=i]} \ = \ h(T_n) \times 1 \ = \ h(T_n)}$$

- Ensuite :
$$\sum_{i=1}^{n} h(T_n)\, \mathbb{1}_{[J=i]} \ = \ \sum_{i=1}^{n} h(S_n - X_J + X_J^*)\, \mathbb{1}_{[J=i]}$$

  Soit $i \in [\![1,n]\!]$. Démontrons alors : $h(S_n - X_J + X_J^*)\, \mathbb{1}_{[J=i]} = h(S_n - X_i + X_i^*)\, \mathbb{1}_{[J=i]}$.
  Soit $\omega \in \Omega$. Deux cas se présentent :

  × si $\omega \notin [J=i]$, alors : $\mathbb{1}_{[J=i]}(\omega) = 0$. Ainsi :

    - d'une part :
    $$\big(h(S_n - X_J + X_J^*)\big)(\omega)\, \mathbb{1}_{[J=i]}(\omega) = \big(h(S_n - X_J + X_J^*)\big)(\omega) \times 0 = 0$$

    - d'autre part :
    $$\big(h(S_n - X_i + X_i^*)\big)(\omega)\, \mathbb{1}_{[J=i]}(\omega) = \big(h(S_n - X_i + X_i^*)\big)(\omega) \times 0 = 0$$

  × si $\omega \in [J=i]$, alors : $\mathbb{1}_{[J=i]}(\omega) = 1$. Ainsi :
  $$\big(h(S_n - X_J + X_J^*)\big)(\omega)\, \mathbb{1}_{[J=i]}(\omega) \ = \ h\big((S_n - X_J + X_J^*)(\omega)\big) \times 1 \ = \ h\big(S_n(\omega) - X_{J(\omega)}(\omega) + X_{J(\omega)}^*(\omega)\big)$$

Or :

- par définition de X_J :

$$\begin{aligned} X_{J(\omega)}(\omega) &= \sum_{j=1}^{n} X_j(\omega)\, \mathbb{1}_{[J=j]}(\omega) \\ &= \sum_{\substack{j=1\\ j\neq i}}^{n} X_j(\omega)\, \mathbb{1}_{[J=j]}(\omega) + X_i(\omega)\, \mathbb{1}_{[J=i]}(\omega) \\ &= \cancel{\sum_{\substack{j=1\\ j\neq i}}^{n} X_j(\omega) \times 0} \; + \; X_i(\omega) \qquad \textit{(car } \omega \in [J=i] \textit{ et } ([J=j])_{j\in[\![1,n]\!]} \textit{ est une famille d'événements incompatibles)} \\ &= X_i(\omega) \end{aligned}$$

- de même, par définition de X_J^* : $X_{J(\omega)}^*(\omega) = X_i^*$.

On en déduit :

$$h\big(S_n(\omega) - X_{J(\omega)}(\omega) + X_{J(\omega)}^*(\omega)\big) = h\big(S_n(\omega) - X_i(\omega) + X_i^*(\omega)\big) = h\big(S_n(\omega) - X_i(\omega) + X_i^*(\omega)\big)\, \mathbb{1}_{[J=i]}(\omega)$$

Et alors :

$$\big(h(S_n - X_J + X_J)\big)(\omega)\, \mathbb{1}_{[J=i]}(\omega) = \big(h(S_n - X_i + X_i^*)\big)(\omega)\, \mathbb{1}_{[J=i]}(\omega)$$

$$\boxed{\text{D'où : } \forall i \in [\![1,n]\!],\ h(S_n - X_J + X_J^*)\, \mathbb{1}_{[J=i]} = h(S_n - X_i + X_i^*)\, \mathbb{1}_{[J=i]}.}$$

- On obtient alors :

$$\sum_{i=1}^{n} h(S_n - X_J + X_J^*)\, \mathbb{1}_{[J=i]} = \sum_{i=1}^{n} h(S_n - X_i + X_i^*)\, \mathbb{1}_{[J=i]}$$

$$\boxed{\text{Finalement : } \sum_{i=1}^{n} h(T_n)\, \mathbb{1}_{[J=i]} = \sum_{i=1}^{n} h(S_n - X_i + X_i^*)\, \mathbb{1}_{[J=i]}.}$$

□

(ii) En déduire : $\mathbb{E}\big(h(T_n)\big) = \sum_{i=1}^{n} \mathbb{P}(\,[J=i]\,)\, \mathbb{E}\big(h(S_n - X_i + X_i^*)\big)$.

Démonstration.

- Tout d'abord, comme la fonction h est bornée, alors $h(T_n)$ admet une espérance, et pour tout $i \in [\![1,n]\!]$, les v.a.r. $h(S_n - X_i + X_i^*)$ admettent une espérance.

Commentaire

- On se permet d'être bref sur l'existence de ces espérances, car c'est sans doute ce qu'attendait l'énoncé à ce stade du sujet.
- Détaillons néanmoins la démonstration ici. On l'effectue dans le cas général. Plus précisément, nous allons démontrer que, pour toute v.a.r. Z quelconque, comme la fonction h est bornée, alors la v.a.r. $h(Z)$ admet une espérance. En effet :
  × Comme h est bornée, alors il existe $M \in \mathbb{R}_+$ tel que : $\forall x \in \mathbb{R}$, $|h(x)| \leqslant M$. Ainsi :

$$0 \leqslant |h(Z)| \leqslant M$$

  × Or la v.a.r. constante égale à M admet une espérance.

  Par théorème d'existence de l'espérance par domination, la v.a.r. $h(Z)$ admet une espérance.

- Soit $i \in [\![1,n]\!]$. On sait que :
  - × les v.a.r. X_1, X_1^*, ..., X_n, X_n^* sont toutes indépendantes de J. Ainsi, par lemme des coalitions, la v.a.r. $h(S_n - X_i + X_i^*)$ est indépendante de la v.a.r. $\mathbb{1}_{[J=i]}$.
  - × la v.a.r. $\mathbb{1}_{[J=i]}$ admet une espérance car c'est une v.a.r. finie.
  - × la v.a.r. $h(S_n - X_i + X_i^*)$ admet une espérance.

$$\boxed{\text{On en déduit que, pour tout } i \in [\![1,n]\!], \text{ la v.a.r. } h(S_n - X_i + X_i^*)\,\mathbb{1}_{[J=i]} \text{ admet une espérance.}}$$

- D'après la question précédente, on obtient :

$$\begin{aligned}
\mathbb{E}\big(h(T_n)\big) &= \mathbb{E}\left(\sum_{i=1}^{n} h(S_n - X_i + X_i^*)\,\mathbb{1}_{[J=i]}\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\Big(h(S_n - X_i + X_i^*)\,\mathbb{1}_{[J=i]}\Big) && \textit{(par linéarité de l'espérance)} \\
&= \sum_{i=1}^{n} \mathbb{E}\Big(h(S_n - X_i + X_i^*)\Big) \times \mathbb{E}\Big(\mathbb{1}_{[J=i]}\Big) && \textit{(car, pour tout } i \in [\![1,n]\!],\ h(S_n - X_i + X_i^*) \textit{ et } \mathbb{1}_{[J=i]} \textit{ sont indépendantes)}
\end{aligned}$$

- Soit $i \in [\![1,n]\!]$. Par définition de $\mathbb{1}_{[J=i]}$:
  - × $\mathbb{1}_{[J=i]}(\Omega) \subset \{0,1\}$
  - × $\mathbb{P}\big(\big[\mathbb{1}_{[J=i]} = 1\big]\big) = \mathbb{P}([J=i])$

$$\boxed{\text{On en déduit, pour tout } i \in [\![1,n]\!] : \mathbb{1}_{[J=i]} \hookrightarrow \mathcal{B}\left(\mathbb{P}([J=i])\right).}$$

- En particulier, pour tout $i \in [\![1,n]\!]$: $\mathbb{E}\Big(\mathbb{1}_{[J=i]}\Big) = \mathbb{P}([J=i])$.

$$\boxed{\text{Finalement : } \mathbb{E}\big(h(T_n)\big) = \sum_{i=1}^{n} \mathbb{E}\big(h(S_n - X_i + X_i^*)\big)\,\mathbb{P}([J=i]).}$$

□

d) Pour $i \in \{1,\dots,n\}$, démontrer, pour tout réel s : $\mathbb{E}\big(h(s + X_i^*)\big) = \dfrac{1}{\mu_i}\,\mathbb{E}\big(X_i\,h(s + X_i)\big)$.

Commentaire

- Notons que, comme la fonction h définie par l'énoncé en question ***9.c)*** est définie sur $[0,+\infty[$, les variables aléatoires $h(s + X_i^*)$ et $h(s + X_i)$ ne sont pas définies pour tout $s \in \mathbb{R}$.
- Pour pallier cette nouvelle inexactitude de l'énoncé, nous considèrerons $s \in \mathbb{R}_+$. Ceci est d'ailleurs sans doute ce que les concepteurs avaient en tête puisque le résultat admis après cette question permet d'identifier le réel s à une réalisation de la v.a.r. $S_n - X_i = \sum\limits_{\substack{j=1 \\ j \neq i}}^{n} X_j$, qui est une v.a.r. positive (en tant que somme de v.a.r. positives).

Démonstration.
Soit $i \in [\![1,n]\!]$. Soit $s \in \mathbb{R}_+$.
On note $\tilde{h}$ la fonction définie par :

$$\tilde{h} : \begin{array}{ccc} [0,+\infty[& \to & \mathbb{R} \\ x & \mapsto & h(s+x) \end{array}$$

*(Notons que la fonction $\tilde{h}$ est bien définie car, comme $s \in \mathbb{R}_+$, alors : $s + x \in [0,+\infty[$.
Et $[0,+\infty[$ est l'ensemble de définition de h.)*

On sait alors que la fonction $\tilde{h} : [0, +\infty[\to \mathbb{R}$ est :

- × bornée, car la fonction h l'est,
- × continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points, car la fonction h l'est.
(Plus précisément, si x_1, …, x_N sont les points de discontinuité de h, alors $x_1 - s$, …, $x_N - s$ sont les points de discontinuité de $\tilde{h}$)

Par définition de la v.a.r. X_i^* qui suit la loi de X_i biaisée par la taille :

$$\begin{array}{ccc} \mathbb{E}\big(\tilde{h}(X_i^*)\big) & = & \dfrac{1}{\mathbb{E}(X_i)}\, \mathbb{E}\big(X_i\, \tilde{h}(X_i)\big) \\ \shortparallel & & \shortparallel \\ \mathbb{E}\big(h(s + X_i^*)\big) & & \dfrac{1}{\mu_i}\, \mathbb{E}\big(X_i\, h(s+X_i)\big) \end{array}$$

$$\boxed{\forall i \in [\![1, n]\!],\ \forall s \in \mathbb{R}_+,\ \mathbb{E}\big(h(s + X_i^*)\big) = \dfrac{1}{\mu_i}\, \mathbb{E}\big(X_i\, h(s+X_i)\big)}$$

□

On admettra qu'on en déduit l'égalité : $\mathbb{E}\big(h(S_n - X_i + X_i^*)\big) = \dfrac{1}{\mu_i}\, \mathbb{E}\big(X_i\, h(S_n)\big)$.

e) En déduire : $\mathbb{E}\big(h(T_n)\big) = \dfrac{\mathbb{E}\big(S_n\, h(S_n)\big)}{\mathbb{E}(S_n)}$.

Démonstration.
D'après la question ***9.c)(ii)*** :

$$\begin{array}{rcll} \mathbb{E}\big(h(T_n)\big) & = & \sum\limits_{i=1}^{n} \mathbb{P}([J = i])\ \mathbb{E}\big(h(S_n - X_i + X_i^*)\big) & \\ & = & \sum\limits_{i=1}^{n} \mathbb{P}([J = i]) \times \dfrac{1}{\mu_i}\, \mathbb{E}\big(X_i\, h(S_n)\big) & \textit{(d'après le résultat admis par l'énoncé)} \\ & = & \sum\limits_{i=1}^{n} \dfrac{\cancel{\mu_i}}{\mu} \times \dfrac{1}{\cancel{\mu_i}}\, \mathbb{E}\big(X_i\, h(S_n)\big) & \textit{(par définition de } J\textit{)} \\ & = & \dfrac{1}{\mu} \sum\limits_{i=1}^{n} \mathbb{E}\big(X_i\, h(S_n)\big) & \\ & = & \dfrac{1}{\mathbb{E}(S_n)} \sum\limits_{i=1}^{n} \mathbb{E}\big(X_i\, h(S_n)\big) & \textit{(d'après } \textbf{9.a)}\textit{)} \\ & = & \dfrac{1}{\mathbb{E}(S_n)}\, \mathbb{E}\left(\sum\limits_{i=1}^{n} X_i\, h(S_n)\right) & \textit{(par linéarité de l'espérance)} \\ & = & \dfrac{1}{\mathbb{E}(S_n)}\, \mathbb{E}\left(h(S_n) \sum\limits_{i=1}^{n} X_i\right) & \\ & = & \dfrac{1}{\mathbb{E}(S_n)}\, \mathbb{E}\big(h(S_n)\, S_n\big) & \end{array}$$

$$\boxed{\mathbb{E}\big(h(T_n)\big) = \dfrac{\mathbb{E}\big(S_n\, h(S_n)\big)}{\mathbb{E}(S_n)}}$$

□

f) Conclure que T_n suit la loi de S_n biaisée par la taille.

Démonstration.
La v.a.r. $S_n = \sum\limits_{i=1}^{n} X_i$ est une variable aléatoire réelle :

× positive, en tant que somme de v.a.r. positives,

× admettant une espérance strictement positive. En effet, d'après ***9.a)***, la v.a.r. S_n admet une espérance et :

$$\mathbb{E}(S_n) \ = \ \sum_{i=1}^{n} \mathbb{E}(X_i) \ > \ 0$$

× telle que, pour toute fonction $h : [0, +\infty[\ \to \mathbb{R}$ bornée et continue sauf éventuellement en un nombre fini de points :

$$\mathbb{E}\big(h(T_n)\big) = \frac{1}{\mathbb{E}(S_n)} \ \mathbb{E}\big(S_n \, h(S_n)\big)$$

Ceci a été démontré avec les questions ***9.c)*** à ***9.e)***.

$$\boxed{\text{D'après la définition fournie avant la question } \textbf{\textit{6.}}, \text{ la v.a.r. } T_n \text{ suit la loi de } S_n \text{ biaisée par la taille.}}$$

□

Troisième partie : Applications en Statistique

On s'intéresse maintenant au cas où le biais par la taille peut être utilisé en statistique, pour construire des estimateurs non biaisés. Une compagnie d'électricité possède n clients où n est un entier naturel non nul donné. Lors de l'année écoulée, le $i^{\text{ème}}$ client a payé x_i euros ($x_i > 0$), mais a en réalité consommé une quantité d'électricité correspondant à y_i euros ($y_i > 0$). La compagnie sait combien ses clients ont payé, et elle souhaite estimer le rapport :

$$\theta \ = \ \frac{\sum\limits_{i=1}^{n} y_i}{\sum\limits_{i=1}^{n} x_i}$$

pour déterminer à quel point elle a mal facturé ses clients.

10. Soit m un entier fixé tel que $1 \leqslant m \leqslant n$. On note $\mathcal{P}_m$ l'ensemble des parties $A \subset \{1, \ldots, n\}$ de cardinal m. On considère une variable aléatoire R, à valeurs dans $\mathcal{P}_m$ et de loi uniforme, c'est-à-dire telle que pour toute partie $A \in \mathcal{P}_m$: $\mathbb{P}\big(\ [R = A]\ \big) = \dfrac{1}{\binom{n}{m}}$.

Commentaire

- Remarquons tout d'abord que la variable aléatoire R `N'est PAS` une variable aléatoire **réelle**. En effet :
  - × comme son nom l'indique, une variable aléatoire réelle prend ses valeurs dans $\mathbb{R}$ (une variable aléatoire réelle X peut prendre la valeur 3 ou $\sqrt{2}$ par exemple).
  - × la variable aléatoire R est à valeurs dans un ensemble de parties. Autrement dit, il s'agit d'une variable aléatoire qui peut prendre pour valeur un **ensemble** (la variable aléatoire R peut prendre la valeur $\{1\}$ ou $\varnothing$ par exemple).

  Cette partie commence donc par l'introduction d'un objet hors norme et nulle part défini dans le programme officiel. Ce dernier se limite en effet à l'étude des variables aléatoires réelles. Le fait qu'une v.a.r. prenne ses valeurs dans $\mathbb{R}$ est loin d'être anecdotique. C'est ce qui permet la bonne définition de la notion de fonction de répartition (pour considérer l'événement $[X \leqslant x]$ par exemple, encore faut-il que l'on puisse comparer x aux valeurs que peut prendre X !), de la notion d'espérance, de la notion de variance. C'est ce qui permet de considérer des transformées affines de v.a.r. , ou encore la somme de v.a.r. , ou encore un produit de v.a.r. . C'est ce qui permet de considérer une transformée $h(X)$ (où h est une fonction $h : X(\Omega) \to \mathbb{R}$). Par exemple, s'il est simple de considérer e^X pour X un variable aléatoire à valeurs réelles, on ne peut donner de sens à e^R puisque R prend pour valeurs des ensembles !
- Pour s'assurer de la bonne compréhension de la variable aléatoire R, commençons par faire un point sur la notion de partie d'un ensemble. Pour fixer les idées, on peut considérer $n = 3$ et noter $E = \{1, 2, 3\}$. Une partie de E n'est autre qu'un sous-ensemble de E. Ainsi, l'**ensemble** des parties de E, noté $\mathcal{P}(E)$, est un **ensemble** d'ensembles défini comme suit :

$$\begin{array}{rclll} \mathcal{P}(E) & = & \{ & \varnothing, & \textit{(la partie vide)} \\ & & & \{1\}, \{2\}, \{3\}, & \textit{(les parties à 1 élément)} \\ & & & \{1,2\}, \{1,3\}, \{2,3\}, & \textit{(les parties à 2 éléments)} \\ & & & \{1,2,3\} \qquad \qquad \} & \textit{(la partie pleine)} \end{array}$$

  L'ensemble $\mathcal{P}_m$ est construit en ne sélectionnant que les parties de cardinal m. On a donc :

$$\mathcal{P}_1 = \Big\{\{1\}, \{2\}, \{3\}\Big\}, \quad \mathcal{P}_2 = \Big\{\{1,2\}, \{1,3\}, \{2,3\}\Big\}, \quad \mathcal{P}_3 = \Big\{\{1,2,3\}\Big\}$$

Commentaire

- Continuons les précisions précédentes en fixant $m = 2$ par exemple. La variable aléatoire R prend alors ses valeurs dans $\mathcal{P}_2 = \Big\{\{1,2\},\{1,3\},\{2,3\}\Big\}$, ensemble qui contient 3 ensembles ($\text{Card}\,\big(\mathcal{P}_2\big) = 3$). Il est précisé que R est de loi uniforme. Ainsi, R prend pour valeur chacun de ces ensembles avec la même probabilité. Plus précisément :
$$\mathbb{P}\big([R = \{1,2\}]\big) \ = \ \mathbb{P}\big([R = \{1,3\}]\big) \ = \ \mathbb{P}\big([R = \{2,3\}]\big) \ = \ \frac{1}{3}$$
- Remarquons enfin que le coefficient binomial $\binom{n}{m}$, n'est autre, par définition, que le nombre de parties à m éléments d'un ensemble à n éléments. Autrement dit, on a : $\text{Card}\,\big(\mathcal{P}_m\big) = \binom{n}{m}$. La variable aléatoire R prenant pour valeur chacun des ensembles de $\mathcal{P}_m$ avec la même probabilité, on a bien :
$$\forall A \in \mathcal{P}_m, \mathbb{P}\big(\ [R = A]\ \big) \ = \ \frac{1}{\binom{n}{m}}$$
- D'après le point précédent, la variable aléatoire R prend un nombre fini de valeurs. Si on conserve l'exemple $n = 3$ et $m = 2$, l'ensemble $\mathcal{P}_2$ est fini et de cardinal : $\text{Card}\,\big(\mathcal{P}_2\big) = \binom{3}{2} = 3$. Cet ensemble peut donc être mis en bijection avec l'ensemble $[\![1,3]\!]$. Une telle bijection revient à numéroter les ensembles de $\mathcal{P}_2$. Par exemple, on peut numéroter comme suit :
$$\begin{array}{ccc} \{1,2\} & \longleftrightarrow & 1 \\ \{1,3\} & \longleftrightarrow & 2 \\ \{2,3\} & \longleftrightarrow & 3 \end{array}$$
À l'aide d'une telle bijection (c'est-à-dire une numérotation des ensembles de $\mathcal{P}_2$), chaque ensemble de $\mathcal{P}_2$ est identifié par un numéro. On aurait pu se servir de cet identifiant pour définir R. Plus précisément, au lieu de définir la variable aléatoire R qui prend pour valeur un ensemble de $\mathcal{P}_2$, on peut définir la variable aléatoire $\tilde{R}$ qui prend pour valeur l'identifiant d'un ensemble de $\mathcal{P}_2$. La variable aléatoire $\tilde{R}$ ainsi définie est une variable aléatoire **réelle**, d'ensemble image $\tilde{R}(\Omega) = [\![1,3]\!]$ et telle que :
$$[R = \{1,2\}] = \Big[\tilde{R} = 1\Big], \quad [R = \{1,3\}] = \Big[\tilde{R} = 2\Big], \quad [R = \{2,3\}] = \Big[\tilde{R} = 3\Big]$$
On peut généraliser cette idée. En considérant une bijection φ de $\mathcal{P}_m$ dans $[\![1,\binom{n}{m}]\!]$, on peut définir $\tilde{R}$ telle que :
$$\forall A \in \mathcal{P}_m, [R = A] = \Big[\tilde{R} = \varphi(A)\Big]$$
- On peut alors légitimement s'interroger sur la pertinence d'introduire la variable aléatoire R en lieu et place de la variable aléatoire **réelle** $\tilde{R}$ qui a l'énorme avantage d'être un objet qui entre dans le cadre du programme officiel. Le pari du concepteur est certainement que le coût de définition de $\tilde{R}$ (en l'occurrence, l'introduction d'une bijection φ de $\mathcal{P}_m$ dans $[\![1,\binom{n}{m}]\!]$) est trop élevé d'un point de vue conceptuel et risque plus de destabiliser les candidats que de les aider. Ainsi, le choix fait dans l'énoncé peut s'entendre. Il est d'ailleurs particulièrement intéressant de voir comment un candidat s'en tire, en fin d'énoncé, avec une notion jamais vue en cours. Cependant, si ce choix est légitime, il eût été souhaitable qu'il s'accompagnât d'une brève mise en garde des candidats pour souligner l'introduction d'une nouvelle notion (celle de variable prenant pour valeurs des ensembles). Ce faisant, on peut considérer que le coût d'introduction de R se rapproche de celui de $\tilde{R}$. L'impression laissée par ce choix est donc que le concepteur préfère « mettre la poussière sous le tapis ». Dans ce cas, on peut supposer que les correcteurs auront pour consigne de ne pas être trop regardant et auront tendance à attribuer la majeure partie des points même en cas de confusions d'objets.

On souhaite écrire un programme pour choisir l'ensemble R au hasard.

a) On considère la procédure suivante : on prend un premier élément s_1 uniformément dans $\{1,\ldots,n\}$, puis un deuxième élément s_2 uniformément dans $\{1,\ldots,n\}\setminus\{s_1\}$, etc. puis un $m^{\text{ème}}$ élément s_m uniformément dans $\{1,\ldots,n\}\setminus\{s_1,\ldots,s_{m-1}\}$. On note $S=(s_1,\ldots,s_m)$, qui est un m-uplet aléatoire.

Commentaire

- Là encore, la présentation peut interroger. On introduit la notion de « m-uplet aléatoire » au lieu de signaler que S est une variable aléatoire qui prend pour valeur le m-uplet construit par le procédé décrit. Ce faisant, l'énoncé utilise la notation $S=(s_1,\ldots,s_m)$ qui porte à confusion. Pour plus de simplicité, illustrons ce type de confusion pour une v.a.r. X :
  - × écrire $X=1$ signifie que la v.a.r. X est constante égale à 1,
  - × dire que X prend la valeur 1 signifie que l'événement $[X=1]$ est réalisé.

  Le concepteur assumant une telle confusion, on peut encore s'attendre à ce que les correcteurs n'aient pas de consigne de fermeté concernant les confusions d'objets.
- On peut par ailleurs s'étonner du manque d'homogénéité de l'énoncé : si on définit S comme dans l'énoncé, on peut tout aussi bien opter pour une présentation similaire pour R. Il s'agit alors de signaler que R est un ensemble aléatoire obtenu par tirage simultané de m nombres de $[\![1,n]\!]$. La question ***10.a)(ii)*** démontre que c'est plutôt cette dernière présentation qui s'inscrit dans l'esprit de l'énoncé.
- Pour mieux appréhender la procédure de choix du m-uplet décrit dans l'énoncé, il suffit de songer à une urne qui contient n-boules numérotées de 1 à n.
  On définit alors les variables aléatoires R et S comme suit :
  - × la variable aléatoire R prend pour valeur le résultat d'un tirage simultané de m boules dans l'urne,
  - × la variable aléatoire S prend pour valeur le résultat de tirages successifs et sans remise de m boules dans l'urne.

(i) Montrer que pour tout m-uplet $(a_1,\ldots,a_m)$ d'entiers distincts de $\{1,\ldots,n\}$, on a :

$$\mathbb{P}(\,[S=(a_1,\ldots,a_m)]\,) \;=\; \frac{(n-m)!}{n!}$$

Démonstration.
Pour tout $i\in[\![1,m]\!]$, notons X_i la v.a.r. qui donne le résultat du $i^{\text{ème}}$ tirage de la procédure décrite dans l'énoncé.

- Remarquons tout d'abord :

$$[S=(a_1,\ldots,a_m)] \;=\; [X_1=a_1]\ \cap\ [X_2=a_2]\ \cap\ \ldots\ \cap\ [X_m=a_m]$$

- Ainsi, par la formule des probabilités composées :

$$\begin{aligned}
&\mathbb{P}\big(\,[S=(a_1,\ldots,a_m)]\,\big)\\
=\;&\mathbb{P}\big(\,[X_1=a_1]\,\big)\times\mathbb{P}_{[X_1=a_1]}\big(\,[X_2=a_2]\,\big)\times\ldots\times\mathbb{P}_{[X_1=a_1]\cap\ldots\cap[X_{m-1}=a_{m-1}]}\big(\,[X_m=a_m]\,\big) \qquad (*)\\
=\;&\frac{1}{n}\times\frac{1}{n-1}\times\ldots\times\frac{1}{n-m+1}
\end{aligned}$$

(car chaque tirage est effectué de manière uniforme dans un ensemble initialement à n éléments et dont le cardinal décrémente de 1 à chaque étape)

L'égalité $(*)$ est justifiée par le fait que : $\mathbb{P}\big(\,[X_1=a_1]\cap\ldots\cap[X_{m-1}=a_{m-1}]\,\big)\neq 0$.

- Or, on remarque :

$$\begin{aligned}
& \frac{1}{n} \times \frac{1}{n-1} \times \ldots \times \frac{1}{n-m+1} \\
= \ & \frac{1}{n \times (n-1) \times \ldots \times (n-m+1)} \\
= \ & \frac{(n-m) \times (n-(m-1)) \times \ldots \times 1}{n \times (n-1) \times \ldots \times (n-m+1) \times (n-m) \times (n-(m-1)) \times \ldots \times 1} \ = \ \frac{(n-m)!}{n!}
\end{aligned}$$

Finalement, on a bien, pour tout m-uplet $(a_1, \ldots, a_m)$ d'entiers distincts de $\{1, \ldots, n\}$:
$$\mathbb{P}(\,[S = (a_1, \ldots, a_m)]\,) \ = \ \frac{(n-m)!}{n!}.$$

Commentaire

- Comme cela a déjà été signalé dans les précédentes remarques, le concepteur ne fait pas de choix affirmé concernant les objets manipulés. En particulier, on ne met pas en avant que S est une variable aléatoire. Cette demi-mesure ne permet pas de comprendre clairement l'esprit du sujet. Globalement, le concepteur pouvait faire deux choix différents :

1) mettre en avant que S, comme R, sont des variables aléatoires. Il est alors naturel, pour résoudre cette question, d'introduire les v.a.r. X_i.

2) insister uniquement sur la procédure d'obtention des m-uplets (en introduisant éventuellement une urne à n boules numérotées de 1 à n). On cherche alors combien de m-uplets différents peuvent être produits par la procédure. Il s'agit alors d'une question purement de dénombrement. Plus précisément : combien y'a-t-il de m-uplets distincts d'éléments de l'ensemble $[\![1, n]\!]$?

On a choisi dans la question précédente la première option. Détaillons maintenant la deuxième option.

- Un m-uplet d'éléments distincts de l'ensemble $[\![1, n]\!]$ est entièrement déterminé par :

× le numéro (élément de $[\![1, n]\!]$) en $1^{\text{ère}}$ position : n possibilités.

× le numéro (élément de $[\![1, n]\!]$) en $2^{\text{ème}}$ position, différent du numéro précédent : $n-1$ possibilités.

× ...

× le numéro (élément de $[\![1, n]\!]$) en $m^{\text{ème}}$ position, différent des numéros précédents : $n-m+1$ possibilités.

Il y a donc $n \times (n-1) \times \ldots \times (n-m+1) = \dfrac{n!}{(n-m)!}$ tels m-uplets.

- Cette deuxième option semble être celle qui s'inscrit le plus dans l'esprit de cette question. Pour autant, il est important de souligner que R est une variable aléatoire pour le reste du sujet. Le choix le plus pertinent aurait certainement été d'insister en question ***10*** sur la procédure de tirage et de présenter cette question sous forme d'un dénombrement. L'introduction de la variable aléatoire R peut alors être introduite seulement en question ***11***. □

(ii) On note $R = \{s_1, \ldots, s_m\}$ l'*ensemble* des entiers tirés lors de la procédure décrite plus haut (l'ordre dans lequel ils ont été tirés n'importe plus). Montrer que pour tout ensemble $A = \{a_1, \ldots, a_m\} \subset \{1, \ldots, n\}$ de cardinal m, on a : $\mathbb{P}(\,[R = A]\,) = \dfrac{m!\,(n-m)!}{n!}$.
En déduire que l'ensemble R a été choisi uniformément dans $\mathcal{P}_m$.

Démonstration.
Soit $A = \{a_1, \ldots, a_m\} \subset \{1, \ldots, n\}$.

- Notons $\mathcal{S}_m$ l'ensemble des bijections de l'ensemble $[\![1, m]\!]$ sur $\{a_1, \ldots, a_m\}$ (la réciproque d'une telle bijection n'est autre qu'une numérotation des éléments de $\{a_1, \ldots, a_m\}$).
  On a alors :
  $$[R = \{a_1, \ldots, a_m\}] \ = \ \bigcup_{\sigma \in \mathcal{S}_m} \left(\bigcap_{i=1}^{m} \ [X_i = \sigma(i)] \ \right)$$
  En effet, l'événement $[R = \{a_1, \ldots, a_m\}]$ est réalisé si et seulement si les éléments $a_1, \ldots, a_m$ ont été obtenus, dans un certain ordre, lors de la procédure. Autrement dit, l'événement $[R = \{a_1, \ldots, a_m\}]$ est réalisé si et seulement si il existe une application $\sigma \in \mathcal{S}_m$ (c'est-à-dire un ordre d'obtention) telle que, pour tout $i \in [\![1, m]\!]$, a_i a été obtenu lors du tirage $\sigma^{-1}(a_i)$.
- Et ainsi :
  $$\begin{aligned}
  & \mathbb{P}\big(R = \{a_1, \ldots, a_m\}\big) \\
  = \ & \mathbb{P}\left(\bigcup_{\sigma \in \mathcal{S}_m} \left(\bigcap_{i=1}^{m} \ [X_i = \sigma(i)] \ \right) \right) \\
  = \ & \sum_{\sigma \in \mathcal{S}_m} \mathbb{P}\left(\bigcap_{i=1}^{m} \ [X_i = \sigma(i)] \ \right) && \textit{(car les événements de la famille } \big(\textstyle\bigcap_{i=1}^{m} \ [X_i = \sigma(i)]\,\big)_{\sigma \in \mathcal{S}_m} \textit{ sont 2 à 2 incompatibles)} \\
  = \ & \sum_{\sigma \in \mathcal{S}_m} \mathbb{P}\big(\,[X_1 = \sigma(1)]\,\big) \times \mathbb{P}_{[X_1 = \sigma(1)]}\big(\,[X_2 = \sigma(2)]\,\big) \times \ldots \times \mathbb{P}_{[X_1 = \sigma(1)] \cap \ldots \cap [X_{m-1} = \sigma(m-1)]}\big(\,[X_m = \sigma(m)]\,\big) \ (*) \\
  = \ & \sum_{\sigma \in \mathcal{S}_m} \frac{1}{n} \times \frac{1}{n-1} \times \ldots \times \frac{1}{n-m+1} && \textit{(car les tirages se font de manière uniforme dans un ensemble initialement à } n \textit{ éléments et donc le cardinal décrémente de 1 à chaque étape)} \\
  = \ & \sum_{\sigma \in \mathcal{S}_m} \frac{(n-m)!}{n!} \ = \ m! \times \frac{(n-m)!}{n!} && \textit{(car } \mathrm{Card}\,\big(\mathcal{S}_m\big) = m!\textit{)}
  \end{aligned}$$
  L'égalité $(*)$ est obtenue par formule des probabilités composées comme en question ***10.a)(i)***.
  $$\boxed{\text{Finalement, on a bien : } \forall A \in \mathcal{P}_m, \mathbb{P}\big(\,[R = A]\,\big) = m! \times \frac{(n-m)!}{n!}.}$$
- On en conclut que pour tout $A \in \mathcal{P}_m$, on a :
  $$\mathbb{P}\big(\,[R = A]\,\big) = m! \times \frac{(n-m)!}{n!} = \frac{1}{\frac{n!}{m!\ (n-m)!}} = \frac{1}{\binom{n}{m}}$$

$$\boxed{\text{Comme } \mathrm{Card}\,\big(\mathcal{P}_m\big) = \tbinom{n}{m}\text{, on en conclut que l'ensemble } R \text{ a été choisi uniformément dans } \mathcal{P}_m.}$$

Commentaire

- Là encore, on note un défaut de formalisme de l'énoncé qui commence par introduire une variable aléatoire R (au début de la question ***10***) et la définit autrement plus loin (dans cette question ***10.a)(ii)***. Comme on l'a conclu dans la remarque précédente, le réel objet d'étude de cette question ***10*** est la procédure d'obtention d'un ensemble aléatoire. C'est certainement plus une démonstration par dénombrement qui est attendue ici.
- L'idée derrière cette question est de dénombrer l'ensemble des m-uplets permettant d'obtenir le même ensemble $\{a_1, \dots, a_m\}$. Pour fixer les idées, considérons le cas $n = 9$ et $m = 3$. L'ensemble $\{1, 3, 8\}$ est obtenu si lors des 3 tirages, on a obtenu 1, 3 et 8. L'événement $[S = \{1, 3, 8\}]$ est réalisé par tous les 3-tirages suivants :

  $$(1, 3, 8) \quad (1, 8, 3) \quad (3, 1, 8) \quad (3, 8, 1) \quad (8, 1, 3) \quad (8, 3, 1)$$

  On dénombre ici six 3-uplets différents. Chacun de ces 3-uplets correspond à une manière d'ordonner les éléments 1, 3, 8. Autrement dit, on considère toutes les permutations de l'ensemble $\{1, 3, 8\}$. Il y en a $3! = 6$. Finalement, à chaque parties à 3 éléments de $[\![1, 9]\!]$ correspond $3! = 6$ différents 3-uplets. Il y a donc $3!$ fois moins de parties à 3 éléments de $[\![1, 9]\!]$ que de 3-uplets d'éléments distincts de $[\![1, 9]\!]$.
  De manière générale, il y a $m!$ fois moins de parties à m éléments de $[\![1, n]\!]$ que de m-uplets d'éléments distincts de $[\![1, n]\!]$. Cela permet de conclure :

  $$\binom{n}{m} = \frac{\frac{n!}{(n-m)!}}{m!} = \frac{n!}{m!(n-m)!}$$

- Lors de l'introduction de la variable R, il est précisé : $\forall A \in \mathcal{P}_m, \mathbb{P}\big([R = A]\big) = \dfrac{1}{\binom{n}{m}}$.
  Or, il s'agit de démontrer ici : $\forall A \in \mathcal{P}_m, \mathbb{P}\big([R = A]\big) = m! \times \dfrac{(n-m)!}{n!} = \dfrac{1}{\binom{n}{m}}$.
  La conclusion n'est autre que l'hypothèse faite dans l'énoncé ! Il faut en réalité comprendre que l'on cherche à démontrer que la procédure décrite dans l'énoncé permet de produire un objet (une variable aléatoire R) qui a la propriété attendue d'uniformité. Le problème provient du fait que le concepteur a donné le même nom (à savoir R) à la variable aléatoire qui possède la propriété et celle dont on cherche à démontrer qu'elle possède la propriété. Cette ambiguïté risque, à raison, de troubler le candidat. □

b) Pour un réel x, on note $\lfloor x \rfloor$ sa partie entière, c'est-à-dire le plus grand entier naturel inférieur ou égal à x. Montrer que si U suit la loi uniforme sur $[0, 1[$, alors $X = 1 + \lfloor n\,U \rfloor$ suit la loi uniforme sur $\{1, \dots, n\}$.

Démonstration.

- Comme $U \hookrightarrow \mathcal{U}([0, 1[)$ on a : $U(\Omega) = [0, 1[$.
  On en déduit :

  $$(n\,U)(\Omega) = [0, n[\qquad \text{et} \qquad \big(\lfloor n\,U \rfloor\big)(\Omega) = [\![0, n[\![= [\![0, n-1]\!]$$

  On obtient finalement :

  $$\big(1 + \lfloor n\,U \rfloor\big)(\Omega) = [\![1, n]\!]$$

  $$\boxed{\text{Ainsi : } X(\Omega) = [\![1, n]\!].}$$

- Soit $k \in [\![1, n]\!]$.

$$\begin{aligned}
\mathbb{P}\big([X = k]\big) &= \mathbb{P}\big([1 + \lfloor n\,U \rfloor = k]\big) \\
&= \mathbb{P}\big([\lfloor n\,U \rfloor = k - 1]\big) \\
&= \mathbb{P}\big([k-1 \leqslant n\,U < (k-1)+1]\big) && \textit{(par définition de la partie entière)} \\
&= \mathbb{P}\left(\left[\frac{k-1}{n} \leqslant U < \frac{k}{n}\right]\right) && \textit{(car } n > 0\textit{)} \\
&= F_U\left(\tfrac{k}{n}\right) - F_U\left(\tfrac{k-1}{n}\right) && \textit{(car } U \textit{ est une v.a.r. à densité)} \\
&= \frac{k}{n} - \frac{k-1}{n} && \textit{(car } U \hookrightarrow \mathcal{U}([0,1[)\textit{)}
\end{aligned}$$

- On en conclut :

$$\left\{\begin{array}{l} X(\Omega) = [\![1, n]\!] \\ \forall k \in [\![1, n]\!],\ \mathbb{P}\big([X = k]\big) = \dfrac{1}{n} \end{array}\right.$$

$$\boxed{\text{Ainsi, } X = 1 + \lfloor n\,U \rfloor \text{ suit bien la loi uniforme sur } \{1, \dots, n\}.}$$

Commentaire

- On rappelle qu'on appelle fonction partie entière la fonction suivante.

$$\begin{array}{rccll} \lfloor\,.\,\rfloor & : & \mathbb{R} & \to & \mathbb{Z} \\ & & x & \mapsto & \lfloor x \rfloor = \text{le plus grand entier } m \in \mathbb{Z} \text{ tel que } m \leqslant x \end{array}$$

On peut aussi définir $\lfloor x \rfloor$ comme l'unique entier relatif m vérifiant la propriété :

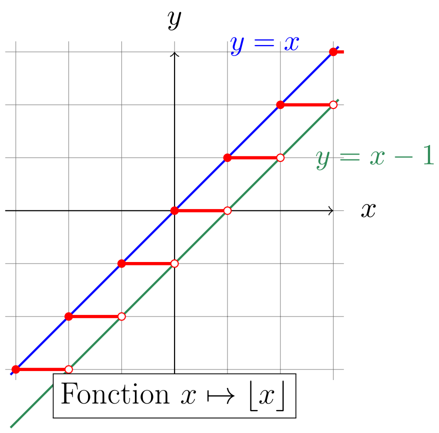
$$m \leqslant x < m + 1$$

- Une propriété fondamentale provenant de la définition de cette fonction est donc :

$$\boxed{\forall u \in \mathbb{R},\ \forall n \in \mathbb{Z},\ \big(\ \lfloor u \rfloor = m \quad \Leftrightarrow \quad m \leqslant u < m + 1\ \big)}$$

En particulier : $\forall x \in \mathbb{R},\ \lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1$.

- Enfin, rappelons que la représentation graphique de la fonction partie entière est la suivante :



Fonction $x \mapsto \lfloor x \rfloor$

□

c) On rappelle que la fonction `rand()` renvoie un nombre aléatoire de loi uniforme sur $[0,1[$, et que `floor(x)` renvoie la partie entière de `x`. Écrire une fonction `Uniforme` en **Scilab** qui prend en argument un entier n, et renvoie un nombre (aléatoire), uniforme sur $\{1,\dots,n\}$.

```
function x = Uniforme(n)
    ...
endfunction
```

Commentaire

- La fonction `rand` permet de simuler une variable aléatoire U de loi uniforme sur $[0,1[$. On préférera cette présentation plutôt que celle de l'énoncé qui se sert de la notion peu rigoureuse de « nombre aléatoire de loi uniforme sur $[0,1[$ ».
- De même, le but de la question est de produire une fonction `Uniforme` qui simule une v.a.r. X telle que $X \hookrightarrow \mathcal{U}(\llbracket 1,n \rrbracket)$.

Démonstration.
Le programme attendu est le suivant.

```
1 function x = Uniforme(n)
2     u = rand()
3     x = 1 + floor(n ⋆ u)
4 endfunction
```

Détaillons les différents éléments de ce programme.

- En ligne 2, on crée tout d'abord la variable `u`.

```
2     u = rand()
```

  Par définition de la fonction `rand`, la variable `u` contient le résultat de la simulation d'une v.a.r. U de loi uniforme sur $[0,1[$.

- Si $U \hookrightarrow \mathcal{U}([0,1[)$, on sait d'après la question précédente que la v.a.r. $X = 1 + \lfloor n\, U \rfloor$ est telle que : $X \hookrightarrow \mathcal{U}(\llbracket 1,n \rrbracket)$. Le résultat attendu par la fonction n'est autre que la simulation de cette v.a.r. X. On le stocke dans la variable `x`.

```
3     x = 1 + floor(n ⋆ u)
```

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 d'obtenir tous les points alloués à cette question.
On procédera de même dans les autres questions **Scilab**. □

d) Écrire une fonction `Selection`, qui prend en argument un vecteur `V` et renvoie un élément `x` de `V` pris de manière aléatoire parmi tous les éléments de `V`, ainsi que le vecteur `W`, égal au vecteur `V` auquel on a enlevé l'élément `x`. L'instruction `length(V)` renvoie le nombre d'éléments du vecteur `V`.

```
function [x, W] = Selection(V)
    n = length(V)
    ...
endfunction
```

Démonstration.
Le programme attendu est le suivant.

```
1 function [x, W] = Selection(V)
2     n = length(V)
3     j = Uniforme(n)
4     x = V(j)
5     W = V( [1:(j-1), (j+1):n] )
6 endfunction
```

Détaillons les différents éléments de ce programme.

- Rappelons que l'instruction `Uniforme(n)` permet de simuler une v.a.r. de loi $\mathcal{U}([\![1, \mathtt{n}]\!])$. Ainsi, en ligne 3, on stocke dans la variable `j` un entier choisi aléatoirement dans $[\![1, n]\!]$.

```
3     j = Uniforme(n)
```

- On stocke alors dans la variable `x` le `j`$^{\text{ème}}$ élément de `V`. C'est le premier résultat attendu de la fonction.

```
4     x = V(j)
```

- Enfin, on stocke dans la variable `W` le contenu du vecteur `V` auquel on a enlevé son `j`$^{\text{ème}}$ élément qui n'est autre que `x`. Rappelons tout d'abord que l'instruction `[1:(j-1), (j+1):n]` pemet de créer une matrice ligne contenant tous les entiers de 1 à `n` à l'exception de l'entier `j`. Ainsi, l'instruction `V( [1:(j-1), (j+1):n] )` permet de sélectionner tous les éléments du vecteur `V` à l'exception du `j`$^{\text{ème}}$ élément.

```
5     W = V( [1:(j-1), (j+1):n] )
```

Commentaire

Dans la question précédente, on a nommé `x` un nombre choisi de manière aléatoire dans l'ensemble $\{1, \dots, \mathtt{n}\}$. Dans cette nouvelle question, `x` ne sert plus à désigner un nombre de $\{1, \dots, \mathtt{n}\}$ mais un élément du vecteur `V` dont on sait seulement qu'il est de taille `n`. Ainsi, d'une question à l'autre, `x` désigne un numéro `j` de $\{1, \dots, \mathtt{n}\}$ puis le `j`$^{\text{ème}}$ élément d'un vecteur. Il était certainement plus pertinent, pour coller à la procédure décrite dans l'énoncé, de stocker le premier résultat de cette fonction dans une variable nommé `a`. □

e) Compléter le programme suivant, qui prend en argument deux entiers n et m avec $m \leqslant n$, et renvoie un vecteur `R` de m entiers distincts, pris uniformément dans $\{1, \ldots, n\}$:

```
function R = Choix(m, n)
    V = 1:n
    R = []
    for i = 1:m
        ...
    end
endfunction
```

Démonstration.
Le programme attendu est le suivant.

```
function R = Choix(m, n)
    V = 1:n
    R = []
    for i = 1:m
        [x, V] = Selection(V)
        R = [R, x]
    end
endfunction
```

Détaillons les différents éléments de ce programme.

- **Début du programme**
  - La variable `V` sert à modéliser l'ensemble dans lequel les tirages vont s'effectuer. Initialement, le tirage s'effectue dans l'ensemble $\{1, \ldots, n\}$. C'est pourquoi on initialise le vecteur `V` à `1:n` (instruction qui permet d'obtenir un vecteur constitué de tous les entiers compris entre `1` et `n`). Ce vecteur `V` sera mis à jour lors du programme.

    ```
        V = 1:n
    ```

  - La variable **R** va accumuler, au fur et à mesure, chaque résultat du tirage qu'on simule. C'est pourquoi on initialise ce vecteur à la matrice vide : tant qu'on n'a pas effectué de tirage, **V** ne contient rien.

    ```
        R = []
    ```

- **Structure itérative**
  Il s'agit de simuler `m` tirages successifs, et sans remise, dans l'ensemble $\{\mathtt{1}, \ldots, \mathtt{n}\}$.
  L'idée est la suivante. Supposons qu'au début d'un certain tour de boucle $\mathtt{i} \in [\![\mathtt{1}, \mathtt{m}]\!]$:
  × la variable `V` est le vecteur constitué de tous les entiers de l'ensemble $\{\mathtt{1}, \ldots, \mathtt{n}\}$ privé des $\mathtt{i} - 1$ premiers entiers tirés (en particulier, au début du premier tour de boucle, le vecteur `V` contient tous les entiers de $\{\mathtt{1}, \ldots, \mathtt{n}\}$).
  × la variable `R` contient les entiers tirés lors des $\mathtt{i} - \mathtt{1}$ premiers tirages (en particulier, au début du premier tour de boucle, le vecteur `R` ne contient rien).

On met alors à jour les variables `V` et `R`. Plus précisément :

– on simule un tirage dans l'ensemble des entiers contenus dans `V` et on retire de `V` l'élément `x` obtenu lors de ce tirage.

```
[x, V] = Selection(V)
```

– on ajoute à `R` l'élément `x` tiré.

```
R = [R, x]
```

Plus précisément, on stocke dans `R` la concaténation du contenu actuel du vecteur `R` et de la valeur de la variable `x`.

Commentaire

Procéder par concaténation n'a pas beaucoup de sens dans cette question. En effet, on sait dès le début du programme que le vecteur `R` va contenir `m` éléments. Il convient donc de définir initialement `R` comme un vecteur à `m` éléments. Cela peut se faire à l'aide de la fonction **`zeros`** par exemple.

```
R = zeros(1, m)
```

Le `i`$^{\text{ème}}$ tour de boucle permet alors de définir le contenu du `i`$^{\text{ème}}$ élément de `R`.

```
R(i) = x
```

En agissant ainsi, on est assuré qu'au début du $(i+1)^{\text{ème}}$ tour de boucle, les variables `V` et `R` contiennent respectivement :

× le vecteur constitué de tous les entiers de l'ensemble $\{1, \ldots, n\}$ privé des `i` premiers entiers tirés.

× le vecteur constitué de tous les entiers tirés lors des `i` premiers tirages.

- **Fin du programme**
La propriété évoquée ci-dessus permet de conclure qu'à l'issue de la boucle (à l'issue du `m`$^{\text{ème}}$ tour de boucle), les variables `V` et `R` contiennent respectivement :

× le vecteur constitué de tous les entiers de l'ensemble $\{1, \ldots, n\}$ privé des `m` entiers tirés.

× le vecteur constitué de tous les entiers tirés lors des `m` tirages.

On simule ainsi la procédure décrite dans l'énoncé. La question ***10.a)(ii)*** permet de s'assurer que cette procédure permet d'obtenir un vecteur constitué de `m` entiers distincts, pris uniformément dans $\{1, \ldots, n\}$, ce qui répond à la question.

Commentaire

Afin de démontrer la correction de ce programme, nous avons exhibé un **invariant de boucle**. En démontrant que cette propriété est vraie avant chaque tour de boucle, on peut conclure quant au contenu des variables à l'issue de la boucle. □

11. Pour une partie $A \in \mathcal{P}_m$, on définit :

$$\bar{x}_A = \frac{1}{m} \sum_{i \in A} x_i, \quad \bar{y}_A = \frac{1}{m} \sum_{i \in A} y_i, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

La compagnie décide d'utiliser $\theta_R = \dfrac{\bar{y}_R}{\bar{x}_R}$ comme estimateur de θ.

a) On définit deux variables aléatoires $X = \bar{x}_R = \dfrac{1}{m} \sum\limits_{i \in R} x_i$ et $Y = \bar{y}_R = \dfrac{1}{m} \sum\limits_{i \in R} y_i$, qui correspondent aux montants moyens payés et consommés par les m clients du groupe tiré au hasard.

Commentaire

- On prendra garde à ne pas faire de confusion d'objets malgré les notations de l'énoncé :
  - × $\bar{x}_R$ est une variable aléatoire
  - × $\bar{x}_A$ est un réel (c'est même une réalisation possible de la v.a.r. $X = \bar{x}_R$).
- Revenons sur la définition de la v.a.r. $X = \bar{x}_R$ (les mêmes remarques s'appliquent sur la v.a.r. $Y = \bar{y}_R$). Elle peut s'énoncer comme suit :

  si R prend la valeur A, alors X prend la valeur $\bar{x}_A$.

  Cette définition est valide car la famille $\big([R = A]\big)_{A \in \mathcal{P}_m}$ forme un système complet d'événements. Démontrons ce point.
  - × D'après l'énoncé, la v.a.r. R est à valeurs dans $\mathcal{P}_m$, *i.e.* : $R(\Omega) \subset \mathcal{P}_m$. On en déduit :

  $$\bigcup_{A \in \mathcal{P}_m} [R = A] \ = \ [R \in \mathcal{P}_m] \ = \ \Omega$$

  $$\boxed{\bigcup_{A \in \mathcal{P}_m} [R = A] \ = \ \Omega}$$

  - × Soit $(A, B) \in (\mathcal{P}_m)^2$ tel que : $A \neq B$. Alors :

  $$[R = A] \cap [R = B] \ = \ \varnothing$$

  En effet, une réalisation de la v.a.r. R ne peut être simultanément égale à deux parties à m éléments distinctes.

  $$\boxed{\text{Les événements de la famille } \big([R = A]\big)_{A \in \mathcal{P}_m} \text{ sont donc incompatibles deux à deux.}}$$

- Ainsi, en exploitant la notion de v.a.r. indicatrice définie par l'énoncé en **Partie II**, on peut définir X par l'espression suivante :

  $$X \ = \ \sum_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{1}_{[R=A]}$$

  L'avantage de cette expression est de se ramener de la v.a.r. R (variable aléatoire à valeurs dans un ensemble de parties), à des variables aléatoires $\mathbb{1}_{[R=A]}$ (variables alétoires finies à valeurs dans $\{0, 1\}$). On se replace alors dans le cadre du programme et les théorèmes du cours peuvent être appliqués (sous réserve de vérification de leurs hypothèses respectives évidemment).
- Bien que ce ne soit sans doute pas un attendu des concepteurs, nous utiliserons régulièrement cette expression dans la suite du sujet afin de pouvoir justifier de l'existence des objets manipulés et effectuer les calculs demandés.

Commentaire

Enfin, démontrons l'égalité : $X = \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{1}_{[R=A]}$.

Soit $\omega \in \Omega$.

- Comme $\big([R=A]\big)_{A \in \mathcal{P}_m}$ forme un système complet d'événements, alors : $\Omega = \bigcup\limits_{A \in \mathcal{P}_m} [R=A]$.

  Ainsi : $\omega \in \bigcup\limits_{A \in \mathcal{P}_m} [R=A]$. Il existe donc $A_0 \in \mathcal{P}_m$ tel que : $\omega \in [R=A_0]$.
- On en déduit :
  - × d'une part, comme $R(\omega) = A_0$, par définition de X :
  $$X(\omega) \ = \ \bar{x}_{A_0}$$
  - × d'autre part, pour tout $A \in \mathcal{P}_m$ tel que $A \neq A_0$, les événements $[R=A]$ et $[R=A_0]$ sont incompatibles. Ainsi : $\omega \notin [R=A]$. D'où :
  $$\mathbb{1}_{[R=A_0]}(\omega) = 1 \qquad \text{et} \qquad \forall A \in \mathcal{P}_m \text{ tel que } A \neq A_0, \ \mathbb{1}_{[R=A]}(\omega) = 0$$

  On obtient :
  $$\begin{array}{rcl} \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \mathbb{1}_{[R=A]}(\omega) & = & \sum\limits_{\substack{A \in \mathcal{P}_m \\ A \neq A_0}} \bar{x}_A \, \mathbb{1}_{[R=A]}(\omega) \ + \ \bar{x}_{A_0} \, \mathbb{1}_{[R=A_0]}(\omega) \\ & = & \cancel{\sum\limits_{\substack{A \in \mathcal{P}_m \\ A \neq A_0}} \bar{x}_A \times 0} \ + \ \bar{x}_{A_0} \times 1 \\ & = & \bar{x}_{A_0} \end{array}$$

On a bien démontré l'égalité voulue.

(i) Démontrer : $\mathbb{E}(X) = \dbinom{n}{m}^{-1} \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A$.

Démonstration.

- Tout d'abord : $X = \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{1}_{[R=A]}$. Or, pour tout $A \in \mathcal{P}_m$, la v.a.r. $\mathbb{1}_{[R=A]}$ admet une espérance car elle est finie.

$$\boxed{\text{La v.a.r. } X \text{ admet donc une espérance en tant que combinaison linéaire de v.a.r. qui en admettent une.}}$$

- De plus :

$$\begin{array}{rcll} \mathbb{E}(X) & = & \mathbb{E}\left(\sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{1}_{[R=A]}\right) & \\ & = & \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{E}\left(\mathbb{1}_{[R=A]}\right) & \textit{(par linéarité de l'espérance)} \\ & = & \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{P}([R=A]) & \textit{(car, pour tout } A \in \mathcal{P}_m, \ \mathbb{1}_{[R=A]} \hookrightarrow \mathcal{B}\left(\mathbb{P}([R=A])\right)\textit{)} \\ & = & \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A \dfrac{1}{\binom{n}{m}} & \end{array}$$

$$\boxed{\text{On en déduit : } \mathbb{E}(X) = \frac{1}{\binom{n}{m}} \sum\limits_{A \in \mathcal{P}_m} \bar{x}_A.}$$

Commentaire

- Comme précisé dans une remarque précédente, l'énoncé n'attendait sûrement pas l'écriture de la v.a.r. X comme une somme de v.a.r. indicatrices. La réponse espérée à cette question est donc sans doute la suivante.
- Par théorème de transfert :

$$\mathbb{E}(X) \ = \ \mathbb{E}(\bar{x}_R) \ = \ \sum_{A \in \mathcal{P}_m} \bar{x}_A \, \mathbb{P}([R = A]) \ = \ \frac{1}{\binom{n}{m}} \sum_{A \in \mathcal{P}_m} \bar{x}_A$$

- Cependant, le programme d'ECE ne traitant que des variables aléatoires **réelles**, nous ne disposons donc pas d'un tel théorème de transfert (en effet, la v.a.r. R n'est pas une variable aléatoire réelle mais une variable aléatoire à valeurs dans un ensemble de parties). Rappelons donc ici le seul théorème de transfert à notre disposition pour les v.a.r. discrètes.

  Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.

  Soit X une variable aléatoire **réelle** discrète sur $(\Omega, \mathscr{A})$.

  Soit $g : X(\Omega) \to \mathbb{R}$ une application.

  Sous réserve d'existence :

$$\boxed{\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x) \ \mathbb{P}([X = x])}$$

□

(ii) Soit $1 \leqslant i \leqslant n$ un entier naturel. Calculer le nombre de parties $A \in \mathcal{P}_m$ telles que $i \in A$.

Démonstration.
On considère l'ensemble $E = [\![1, n]\!]$ à n éléments.
(on peut penser à une pièce qui contient n individus)
On souhaite construire une partie A à m éléments de cet ensemble contenant un élément distingué : l'élément i *(on peut penser à choisir dans la pièce un groupe de m individus dans lequel figure l'individu numéro i).*
Pour ce faire, on choisit d'abord, dans E, l'élément numéro i : 1 possibilité.
Pour former la partie A, on choisit ensuite $m-1$ éléments dans E, en y ajoutant l'élément i : $\binom{n-1}{m-1}$ possibilités.
(on choisit d'abord l'individu numéro i puis on lui adjoint un groupe de $m-1$ individus)

$$\boxed{\text{Ainsi, il y a } 1 \times \binom{n-1}{m-1} = \binom{n-1}{m-1} \text{ manières de construire une partie } A \in \mathcal{P}_m \text{ telle que } i \in A.}$$

□

(iii) En déduire :

$$\sum_{A \in \mathcal{P}_m} \sum_{i \in A} x_i \ = \ \binom{n-1}{m-1} \sum_{i=1}^{n} x_i$$

Démonstration.

- Tout d'abord :

$$\sum_{A \in \mathcal{P}_m} \left(\sum_{i \in A} x_i \right) \ = \ \sum_{A \in \mathcal{P}_m} \Big(\sum_{\substack{i = 1 \\ i \in A}}^{n} x_i \Big)$$

- On en déduit :

$$\begin{array}{rcll}
\displaystyle\sum_{A\in\mathcal{P}_m}\left(\sum_{i\in A} x_i\right) & = & \displaystyle\sum_{\substack{A\in\mathcal{P}_m,\, 1\leqslant i\leqslant n \\ i\in A}} x_i & \\
& = & \displaystyle\sum_{i=1}^{n}\Bigg(\sum_{\substack{A\in\mathcal{P}_m \\ i\in A}} x_i\Bigg) & \\
& = & \displaystyle\sum_{i=1}^{n}\Bigg(x_i \sum_{\substack{A\in\mathcal{P}_m \\ i\in A}} 1\Bigg) & \\
& = & \displaystyle\sum_{i=1}^{n}\Big(x_i\ \mathrm{Card}\left(\{A\in\mathcal{P}_m \mid i\in A\}\right)\Big) & \\
& = & \displaystyle\sum_{i=1}^{n} x_i \binom{n-1}{m-1} & \textit{(d'après la question précédente)}
\end{array}$$

$$\boxed{\text{Finalement : } \displaystyle\sum_{A\in\mathcal{P}_m}\left(\sum_{i\in A} x_i\right) = \binom{n-1}{m-1}\sum_{i=1}^{n} x_i.}$$

□

(iv) Conclure : $\mathbb{E}(X)=\bar{x}$. On **admettra** que de même on a : $\mathbb{E}(Y)=\bar{y}$.

Démonstration.
D'après la question ***11.a)(i)***, on obtient :

$$\begin{array}{rcll}
\mathbb{E}(X) & = & \dfrac{1}{\binom{n}{m}}\displaystyle\sum_{A\in\mathcal{P}_m}\bar{x}_A & \\
& = & \dfrac{1}{\binom{n}{m}}\displaystyle\sum_{A\in\mathcal{P}_m}\left(\frac{1}{m}\sum_{i\in A} x_i\right) & \textit{(par définition de } \bar{x}_A\textit{)} \\
& = & \dfrac{1}{\frac{n}{\cancel{m}}\binom{n-1}{m-1}}\,\dfrac{1}{\cancel{m}}\left(\displaystyle\sum_{A\in\mathcal{P}_m}\left(\sum_{i\in A} x_i\right)\right) & \textit{(car : } m\binom{n}{m} = n\binom{n-1}{m-1}\textit{)} \\
& = & \dfrac{1}{n\,\cancel{\binom{n-1}{m-1}}}\,\cancel{\binom{n-1}{m-1}}\displaystyle\sum_{i=1}^{n} x_i & \textit{(d'après la question précédente)} \\
& = & \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i &
\end{array}$$

$$\boxed{\text{On en déduit : } \mathbb{E}(X)=\bar{x}.}$$

□

(v) Exprimer θ en fonction de $\mathbb{E}(X)$ et $\mathbb{E}(Y)$.

Démonstration.

Par définition de θ, on obtient : $\theta = \dfrac{\sum\limits_{i=1}^{n} y_i}{\sum\limits_{i=1}^{n} x_i} = \dfrac{\frac{1}{n}\sum\limits_{i=1}^{n} y_i}{\frac{1}{n}\sum\limits_{i=1}^{n} x_i} = \dfrac{\bar{y}}{\bar{x}}$.

$$\boxed{\text{D'après la question précédente, on obtient : } \theta = \frac{\mathbb{E}(Y)}{\mathbb{E}(X)}.}$$

□

b) Démontrer : $\mathbb{E}(\theta_R) = \mathbb{E}\left(\dfrac{Y}{X}\right)$.

Démonstration.
Par définition des v.a.r. X et Y, on sait :

$$X(\Omega) \subset \{\bar{x}_A \mid A \in \mathcal{P}_m\} \qquad \text{et} \qquad Y(\Omega) \subset \{\bar{y}_A \mid A \in \mathcal{P}_m\}$$

Comme l'ensemble $\mathcal{P}_m$ est une ensemble fini, alors les v.a.r. X et Y sont finies.

$$\boxed{\text{La v.a.r. } \dfrac{Y}{X} \text{ admet donc une espérance car c'est une v.a.r. finie.}}$$

$$\boxed{\text{Par définition de } \theta_R\text{, on a de plus : } \mathbb{E}(\theta_R) = \mathbb{E}\left(\dfrac{\bar{y}_R}{\bar{x}_R}\right) = \mathbb{E}\left(\dfrac{Y}{X}\right).}$$

□

c) On donne l'inégalité de Cauchy-Schwarz : si W et Z sont deux variables aléatoires strictement positives, admettant un moment d'ordre deux : $\mathbb{E}(W\,Z) \leqslant \big(\mathbb{E}(W^2)\big)^{\frac{1}{2}} \big(\mathbb{E}(Z^2)\big)^{\frac{1}{2}}$, avec égalité si et seulement s'il existe un $\alpha > 0$ tel que $W = \alpha\, Z$.

Commentaire

- Le cas d'égalité donné par l'énoncé est erroné. En effet :

$$\mathbb{E}(W\,Z) \leqslant \big(\mathbb{E}(W^2)\big)^{\frac{1}{2}} \big(\mathbb{E}(Z^2)\big)^{\frac{1}{2}} \quad \Leftrightarrow \quad \begin{array}{c}\text{il existe } \alpha > 0 \text{ tel que } W = \alpha\, Z \\ \textbf{presque sûrement}\end{array}$$

- Démontrons l'inégalité de Cauchy-Schwarz fournie par l'énoncé et son cas d'égalité.
  - × On note f la fonction polynomiale définie par :

$$f : \lambda \mapsto \mathbb{E}(W^2)\,\lambda^2 + 2\,\mathbb{E}(W\,Z)\,\lambda + \mathbb{E}(Z^2)$$

    La fonction f est une fonction polynomiale de degré 2 en λ.
    On note P le polynôme de degré 2 associé.
  - × Par linéarité de l'espérance :

$$\forall \lambda \in \mathbb{R},\ f(\lambda) = \mathbb{E}\big((\lambda\, W + Z)^2\big)$$

    Par positivité de l'espérance : $\forall \lambda \in \mathbb{R},\ f(\lambda) \geqslant 0$.
  - × La fonction f étant positive, le polynôme associé P est de signe constant.
    Son discriminant Δ est donc négatif ou nul. Or :

$$\Delta \ = \ \big(2\,\mathbb{E}(W\,Z)\big)^2 - 4\,\mathbb{E}(W^2)\,\mathbb{E}(Z^2) \ = \ 4\left(\big(\mathbb{E}(W\,Z)\big)^2 - \mathbb{E}(W^2)\,\mathbb{E}(Z^2)\right)$$

    On en déduit :

$$\begin{array}{ll} \text{comme} & \Delta \leqslant 0 \\ \text{alors} & 4\left(\big(\mathbb{E}(W\,Z)\big)^2 - \mathbb{E}(W^2)\,\mathbb{E}(Z^2)\right) \leqslant 0 \\ \text{donc} & \big(\mathbb{E}(W\,Z)\big)^2 - \mathbb{E}(W^2)\,\mathbb{E}(Z^2) \leqslant 0 \end{array}$$

$$\boxed{\text{D'où : } \big(\mathbb{E}(W\,Z)\big)^2 \leqslant \mathbb{E}(W^2)\,\mathbb{E}(Z^2).}$$

Commentaire

× Déteminons maintenant sous quelle condition a lieu l'égalité de l'énoncé.

(⇒) Supposons : $\big(\mathbb{E}(W\,Z)\big)^2 = \mathbb{E}(W^2)\,\mathbb{E}(Z^2)$.

$$\begin{array}{rl} \text{Alors} & \big(\mathbb{E}(W\,Z)\big)^2 - \mathbb{E}(W^2)\,\mathbb{E}(Z^2) = 0 \\ \text{donc} & 4\,\big(\mathbb{E}(W\,Z)\big)^2 - \mathbb{E}(W^2)\,\mathbb{E}(Z^2) = 0 \\ \text{d'où} & \Delta = 0 \end{array}$$

On en déduit que le polynôme P admet une unique racine. Il existe donc $\lambda_0 \in \mathbb{R}$ tel que :

$$\begin{array}{rll} & P(\lambda_0) = 0 & \\ \text{ainsi} & \mathbb{E}\big((\lambda_0\,W + Z)^2\big) = 0 & \\ \text{d'où} & (\lambda_0\,W + Z)^2 = 0 \text{ presque sûrement} & \textit{(car } (\lambda_0\,W+Z)^2 \textit{ est une v.a.r. positive)} \\ \text{alors} & \lambda_0\,W + Z = 0 \text{ presque sûrement} & \\ \text{enfin} & Z = -\lambda_0\,W \text{ presque sûrement} & \end{array}$$

On note $\alpha = -\lambda_0$. Comme les v.a.r. W et Z sont strictement positives, on en déduit que le réel α est strictement positif.

$$\boxed{\text{On a démontré : } \big(\mathbb{E}(W\,Z)\big)^2 = \mathbb{E}(W^2)\,\mathbb{E}(Z^2) \;\Rightarrow\; \begin{array}{l}\text{il existe } \alpha > 0 \text{ tel que} \\ Z = \alpha\,W, \text{ presque sûrement}\end{array}}$$

(⇐) Réciproquement, supposons qu'il existe $\alpha > 0$ tel que $Z = \alpha\,W$, presque sûrement. Ainsi :

× d'une part :

$$\begin{array}{rcll} \big(\mathbb{E}(W\,Z)\big)^2 & = & \Big(\mathbb{E}\big(W \times (\alpha\,W)\big)\Big)^2 & \\ & = & \big(\alpha\,\mathbb{E}(W^2)\big)^2 & \textit{(par linéarité de l'espérance)} \\ & = & \alpha^2\,\big(\mathbb{E}(W^2)\big)^2 & \end{array}$$

× d'autre part :

$$\mathbb{E}(W^2)\,\mathbb{E}(Z^2) \;=\; \mathbb{E}(W^2)\,\mathbb{E}\big((\alpha\,W)^2\big) \;=\; \mathbb{E}(W^2) \times \alpha^2\,\mathbb{E}(W^2) \;=\; \alpha^2\,\big(\mathbb{E}(W^2)\big)^2$$

On en déduit bien : $\big(\mathbb{E}(W\,Z)\big)^2 = \mathbb{E}(W^2)\,\mathbb{E}(Z^2)$.

$$\boxed{\text{On a démontré : } \begin{array}{l}\text{il existe } \alpha > 0 \text{ tel que} \\ Z = \alpha\,W, \text{ presque sûrement}\end{array} \;\Rightarrow\; \big(\mathbb{E}(W\,Z)\big)^2 = \mathbb{E}(W^2)\,\mathbb{E}(Z^2)}$$

(i) Démontrer : $\mathbb{E}\left(\dfrac{1}{X}\right) \geqslant \dfrac{1}{\mathbb{E}(X)}$.

Démonstration.

- Tout d'abord :

$$\mathbb{E}\left(\frac{1}{X}\right) \geqslant \frac{1}{\mathbb{E}(X)} \quad \Leftrightarrow \quad \mathbb{E}(X)\,\mathbb{E}\left(\frac{1}{X}\right) \geqslant 1 \qquad (\textit{car } \mathbb{E}(X) > 0)$$

- On souhaite alors appliquer l'inégalité de Cauchy-Schwarz à $W = \sqrt{X}$ et $Z = \dfrac{1}{\sqrt{X}}$.
  Ces deux v.a.r. :
  × sont strictement positives, car la v.a.r. X l'est. En effet, d'après l'énoncé : $\forall i \in [\![1,n]\!]$, $x_i > 0$.
  Ainsi, pour tout $\omega \in \Omega$: $X(\omega) = \dfrac{1}{m} \sum\limits_{i \in R(\omega)} x_i > 0$.
  × admettent un moment d'ordre 2, car elles sont finies.
  Par inégalité de Cauchy-Schwarz, on obtient :

$$\begin{array}{ccccc} & & \mathbb{E}\left(\sqrt{X} \times \dfrac{1}{\sqrt{X}}\right) & \leqslant & \sqrt{\mathbb{E}\left((\sqrt{X})^2\right) \times \mathbb{E}\left(\left(\dfrac{1}{\sqrt{X}}\right)^2\right)} \\ & & \shortparallel & & \shortparallel \\ 1 & = & \mathbb{E}(1) & & \sqrt{\mathbb{E}(X) \times \mathbb{E}\left(\dfrac{1}{X}\right)} \end{array}$$

- Par croissance de la fonction $x \mapsto x^2$ sur $[0, +\infty[$, on obtient :

$$1 \leqslant \mathbb{E}(X) \times \mathbb{E}\left(\frac{1}{X}\right)$$

$$\boxed{\text{Avec le premier point de cette démonstration, on en déduit : } \mathbb{E}\left(\frac{1}{X}\right) \geqslant \frac{1}{\mathbb{E}(X)}.}$$

□

(ii) Montrer qu'il y a égalité si et seulement si X est une variable aléatoire constante, c'est-à-dire $X = \mathbb{E}(X) = \bar{x}$.

Démonstration.
Il a déjà été précisé en question précédente que les v.a.r. $W = \sqrt{X}$ et $Z = \dfrac{1}{\sqrt{X}}$ vérifiaient les conditions d'application de l'inégalité de Cauchy-Schwarz. Ainsi, d'après l'énoncé :

$$\begin{array}{rcl} \mathbb{E}\left(\dfrac{1}{X}\right) = \dfrac{1}{\mathbb{E}(X)} & \Leftrightarrow & \exists \alpha > 0,\ \sqrt{X} = \alpha\, \dfrac{1}{\sqrt{X}},\ \text{presque sûrement} \\ & \Leftrightarrow & \exists \alpha > 0,\ \sqrt{X}\,\sqrt{X} = \alpha,\ \text{presque sûrement} \\ & \Leftrightarrow & \exists \alpha > 0,\ X = \alpha,\ \text{presque sûrement} \\ & \Leftrightarrow & \text{la v.a.r. } X \text{ est constante presque sûrement} \end{array}$$

Enfin, une v.a.r. constante presque sûrement est égale à son espérance, presque sûrement.

$$\boxed{\begin{array}{c}\text{Finalement } \mathbb{E}\left(\dfrac{1}{X}\right) = \dfrac{1}{\mathbb{E}(X)} \text{ si et seulement si } X \text{ est presque sûrement} \\ \text{constante, égale à } \mathbb{E}(X) = \bar{x} \text{ (d'après } \textbf{\textit{11.a)(iv)}}\text{).}\end{array}}$$

□

(iii) Conclure que $\mathbb{E}\left(\dfrac{1}{X}\right) = \dfrac{1}{\mathbb{E}(X)}$ si et seulement si pour tout $i \in [\![1,n]\!]$, $x_i = \bar{x}$.

Démonstration.

- Tout d'abord, d'après la question précédente :

$$\mathbb{E}\left(\frac{1}{X}\right) = \frac{1}{\mathbb{E}(X)} \quad \Leftrightarrow \quad X = \bar{x}$$

- Démontrons alors :

$$X = \bar{x} \quad \Leftrightarrow \quad \forall i \in [\![1,n]\!],\ x_i = \bar{x}$$

($\Leftarrow$) Supposons : $\forall i \in [\![1,n]\!]$, $x_i = \bar{x}$.
Soit $\omega \in \Omega$.
Comme $\big([R = A]\big)_{A \in \mathcal{P}_m}$ forme un système complet d'événements, il existe $A \in \mathcal{P}_m$ tel que : $\omega \in [R = A]$. D'où :

$$X(\omega) \ = \ \bar{x}_A \ = \ \frac{1}{m} \sum_{i \in A} x_i \ = \ \frac{1}{m} \sum_{i \in A} \bar{x} \ = \ \frac{1}{m}\ \mathrm{Card}(A)\, \bar{x} \ = \ \frac{1}{m}\, m\, \bar{x} \ = \ \bar{x}$$

($\Rightarrow$) Supposons : $X = \bar{x}$.

- Soit $A \in \mathcal{P}_m$. Alors il existe $\omega \in \Omega$ tel que : $R(\omega) = A$. Ainsi : $X(\omega) = \bar{x}_A$. D'où : $\bar{x} = \bar{x}_A$. Ainsi :

$$\forall A \in \mathcal{P}_m,\ \bar{x}_A = \bar{x}$$

- Soit $i \in [\![2,n]\!]$. Deux cas se présentent :
  - si $i \in [\![1,m]\!]$, alors on considère les ensembles suivants :

$$A = [\![2, m+1]\!] \qquad \text{et} \qquad B = [\![1, m+1]\!] \setminus \{i\}$$

D'après le point précédent : $\bar{x}_A = \bar{x} = \bar{x}_B$. Alors :

$$\begin{array}{rcccl}
 & \bar{x}_A & = & \bar{x}_B \\
\text{donc} & \displaystyle\sum_{k=2}^{m+1} x_k & = & \displaystyle\sum_{\substack{k=1 \\ k \neq i}}^{m} x_k \\
\text{d'où} & \cancel{\displaystyle\sum_{\substack{k=2 \\ k \neq i}}^{m} x_k} + x_i & = & x_1 + \cancel{\displaystyle\sum_{\substack{k=2 \\ k \neq i}}^{m} x_k} \\
\text{ainsi} & x_i & = & x_1
\end{array}$$

  - si $i \notin [\![1,m]\!]$, alors on considère les ensembles suivants :

$$A = [\![2, m]\!] \cup \{i\} \qquad \text{et} \qquad B = [\![1, m]\!]$$

D'après le point précédent : $\bar{x}_A = \bar{x} = \bar{x}_B$. Alors :

$$\begin{array}{rcccl}
 & \bar{x}_A & = & \bar{x}_B \\
\text{donc} & \displaystyle\sum_{k \in A} x_k & = & \displaystyle\sum_{k=1}^{m} x_k \\
\text{d'où} & \cancel{\displaystyle\sum_{k=2}^{m} x_k} + x_i & = & x_1 + \cancel{\displaystyle\sum_{k=2}^{m} x_k} \\
\text{ainsi} & x_i & = & x_1
\end{array}$$

Finalement : $\forall i \in [\![1,n]\!]$, $x_i = x_1$.

- On en déduit :
$$\bar{x} \;=\; \frac{1}{n} \sum_{i=1}^{n} x_i \;=\; \frac{1}{n} \sum_{i=1}^{n} x_1 \;=\; \frac{1}{n}\, n\, x_1 \;=\; x_1$$
D'où : $\forall i \in [\![1,n]\!]$, $x_i = x_1 = \bar{x}$.

Finalement : $\mathbb{E}\left(\dfrac{1}{X}\right) = \dfrac{1}{\mathbb{E}(X)}$ si et seulement si pour tout $i \in [\![1,n]\!]$, $x_i = \bar{x}$.

Commentaire

- Notons que, pour que les ensembles A et B définis dans le cas $i \in [\![1,m]\!]$ existent, il faut imposer : $m < n$.
- Cette hypothèse est même indispensable puisque, dans le cas $n = m$, l'équivalence qu'il est demandé de démontrer est fausse.
  En effet, si $n = m$, alors le seul ensemble de $\mathcal{P}_m = \mathcal{P}_n$ est l'ensemble $A = [\![1,n]\!]$. On en déduit que la variable aléatoire R ne peut prendre que la valeur A. Ainsi, pour tout $\omega \in \Omega$:
  $$X(\omega) \;=\; \bar{x}_A \;=\; \bar{x}_{[\![1,n]\!]} \;=\; \frac{1}{n} \sum_{i=1}^{n} x_i \;=\; \bar{x}$$
  Si $n = m$, on a donc démontré, peu importe les valeurs de $x_1, \ldots, x_n$, que la v.a.r. X est constante égale à $\bar{x}$.
- Dans la suite de ce problème, on se placera toujours dans le cas : $1 \leqslant m < n$.
- On a détaillé ici rigoureusement le choix des ensembles A et B permettant de mener à bien la démonstration de cette question. Il était possible de choisir bien d'autres ensembles. On peut d'ailleurs penser qu'une explication « avec les mains » du choix de ces 2 ensembles suffisait à obtenir la majorité des points alloués à cette questions. □

d) Si on suppose que X et Y sont indépendantes, montrer que $\mathbb{E}(\theta_R) \geqslant \theta$, avec égalité si et seulement si pour tout $i \in [\![1,n]\!]$, $x_i = \bar{x}$.

Démonstration.

- Comme les v.a.r. X et Y sont indépendantes, alors, par lemme des coalitions, les v.a.r. $\dfrac{1}{X}$ et Y le sont aussi. On obtient :
$$\begin{array}{rcll} \mathbb{E}(\theta_R) & = & \mathbb{E}\left(\dfrac{Y}{X}\right) & \textit{(d'après } \textbf{\textit{11.b)}}\textit{)} \\ & = & \mathbb{E}\left(Y \times \dfrac{1}{X}\right) & \\ & = & \mathbb{E}(Y) \times \mathbb{E}\left(\dfrac{1}{X}\right) & \textit{(car } Y \textit{ et } \dfrac{1}{X} \textit{ sont indépendantes)} \end{array}$$
- De plus, d'après la question ***11.c)(i)*** :
$$\begin{array}{rcccl} & \mathbb{E}\left(\dfrac{1}{X}\right) & \geqslant & \dfrac{1}{\mathbb{E}(X)} & \\ \text{donc} & \mathbb{E}(Y)\, \mathbb{E}\left(\dfrac{1}{X}\right) & \geqslant & \mathbb{E}(Y)\, \dfrac{1}{\mathbb{E}(X)} & \textit{(car } \mathbb{E}(Y) > 0\textit{)} \\ \text{d'où} & \mathbb{E}(\theta_R) & \geqslant & \dfrac{\mathbb{E}(Y)}{\mathbb{E}(X)} & \textit{(d'après le point précédent)} \end{array}$$

D'après la question ***11.a)(v)***, on obtient : $\mathbb{E}(\theta_R) \geqslant \theta$.

- On remarque :

$$\begin{array}{rcll} \mathbb{E}(\theta_R) = \mathbb{E}(\theta) & \Leftrightarrow & \mathbb{E}(Y)\,\mathbb{E}\left(\dfrac{1}{X}\right) = \dfrac{\mathbb{E}(Y)}{\mathbb{E}(X)} & \textit{(d'après les points précédents)} \\ & \Leftrightarrow & \mathbb{E}\left(\dfrac{1}{X}\right) = \mathbb{E}(X) & \textit{(car } \mathbb{E}(Y) \neq 0\textit{)} \\ & \Leftrightarrow & \forall i \in [\![1,n]\!],\ x_i = \bar{x} & \textit{(d'après } \textbf{11.c)}\textit{)} \end{array}$$

$$\boxed{\mathbb{E}(\theta_R) = \mathbb{E}(\theta) \Leftrightarrow \forall i \in [\![1,n]\!],\ x_i = \bar{x}}$$

□

Ainsi, $\mathbb{E}(\theta_R)$ n'est pas forcément égal à θ : on dit alors que θ_R est un estimateur *biaisé* de θ.

12. Ce problème peut être résolu en choisissant les m clients non de manière uniforme comme dans la question ***10.***, mais de manière biaisée par la taille. Par analogie avec la construction de T_n dans la question ***9.***, on commence par choisir une variable aléatoire J à valeurs dans $\{1, 2, \dots, n\}$, dont la loi est donnée par : $\mathbb{P}(\ [J = i]\ ) = \dfrac{x_i}{\sum\limits_{r=1}^{n} x_r}$. Ensuite, étant donné J, on choisit un groupe V de $m-1$ clients parmi les $n-1$ clients différents de J, de manière uniforme. Autrement dit, pour toute partie $A \in \mathcal{P}_m$, et tout $i \in A$, on a :

$$\mathbb{P}_{[J=i]}(\,[V = A \setminus \{i\}]\,) \ = \ \frac{1}{\binom{n-1}{m-1}}$$

Le groupe de clients examiné est alors : $R = V \cup \{J\}$.

Commentaire

Malheureusement, comme à chaque question de ce sujet, il convient de s'interroger sur les objets manipulés et les choix faits par le concepteur.

- Il y a tout d'abord un problème de variable libre / liée. La variable i n'est pas introduite. Si on en croit la présentation, cette variable est libre. S'il y avait eu un effort de rigueur dans l'énoncé, celui-ci aurait été :

$$\forall i \in [\![1,n]\!],\ \mathbb{P}(\ [J = i]\ ) = \frac{x_i}{\sum\limits_{r=1}^{n} x_r}$$

  La variable i est donc liée (par un quantificateur universel). Il eût été appréciable que soit fait un autre choix de variable. Cela aurait en effet permis de distinguer avec la variable i (encore une fois non introduite) utilisée après. On aurait pu écrire :

$$\forall j \in [\![1,n]\!],\ \mathbb{P}(\ [J = j]\ ) = \frac{x_j}{\sum\limits_{r=1}^{n} x_r}$$

Commentaire

- L'énoncé poursuit alors par la définition de V. Elle aussi est problématique.
  Il faut tout d'abord comprendre que c'est une variable aléatoire (ce n'est pas explicitement précisé dans l'énoncé). Il faut alors faire en sorte de traduire l'énoncé par des termes mathématiques. Lorsqu'il est dit « étant donné J », il faut comprendre « lorsque J prend une valeur i donnée ». Ici, le concepteur fait le choix de confondre la v.a.r. J et la valeur prise par celle-ci. C'est une manière de présenter les choses tout à fait usuelle pour qui fait des probabilités à un niveau d'études supérieures avancé. Il est par contre légitime de s'interroger sur sa pertinence dans un sujet de la filière ECE.
  Il faut ensuite comprendre que cette variable aléatoire est à valeurs dans $\mathcal{P}_{m-1}$, ensemble dont les éléments sont eux-mêmes des ensembles. Rappelons, comme signalé en question ***10.***, que cette construction n'est pas officiellement au programme.
- Pour mieux appréhender la définition de J et V, on peut réfléchir à l'aide d'une urne à n boules numérotées de 1 à n et à une expérience en deux temps.
  × On effectue tout d'abord un tirage dans l'urne.
  On définit alors la v.a.r. J qui prend comme valeur le numéro associé à cette boule.
  × On décide alors d'effectuer un tirage simultané de $m-1$ boules dans l'urne.
  On définit la v.a.r. V qui prend comme valeur le résultat de ce tirage.

  Ce dernier tirage simultané étant réalisé de manière uniforme, chacune des parties à $m-1$ éléments est obtenue avec la même probabilité. Ainsi, pour tout $i \in [\![1, n]\!]$, la loi conditionnelle de V relativement à l'événement $[J = i]$, est la loi uniforme sur $\mathcal{P}_{m-1}$. Autrement dit :

$$\forall i \in [\![1, n]\!],\ \forall B \in \mathcal{P}_{m-1},\ \mathbb{P}_{[J=i]}(\,[V = B]\,) \ = \ \frac{1}{\binom{n-1}{m-1}}$$

  ou encore, comme le suggère l'énoncé :

$$\forall i \in [\![1, n]\!],\ \forall A \in \mathcal{P}_{m},\ \mathbb{P}_{[J=i]}(\,[V = A \setminus \{i\}]\,) \ = \ \begin{cases} \dfrac{1}{\binom{n-1}{m-1}} & \text{si } i \in A \\ 0 & \text{si } i \notin A \end{cases}$$

  Rappelons que, par définition, il y a exactement $\binom{n-1}{m-1}$ parties à $m-1$ éléments dans un ensemble de $n-1$ éléments.
- L'énoncé définit enfin R. Il faut aussi comprendre que c'est une variable aléatoire qui n'est pas réelle mais à valeurs dans $\mathcal{P}_m$. Étant données les valeurs prises par les variables aléatoires considérées, on peut tout à fait considérer d'effectuer la réunion présentée dans l'énoncé. Toutefois, on eût aimé qu'une telle construction eût été introduite correctement. La variable aléatoire $V \cup \{J\}$ est correctement définie par :

$$\begin{array}{rcrcl} V \cup \{J\} & : & \Omega & \to & \mathcal{P}_m \\ & & \omega & \mapsto & V(\omega) \cup \{J(\omega)\} \end{array}$$

a) On commence par déterminer $\mathbb{P}(\,[R = A]\,)$, pour $A \in \mathcal{P}_m$ donné.

(i) Démontrer :

$$\mathbb{P}(\,[R = A]\,) \;=\; \sum_{i \in A} \mathbb{P}(\,[J = i]\,)\, \mathbb{P}_{[J=i]}(\,[V = A \setminus \{i\}]\,)$$

Démonstration.

- La famille $\big(\,[J = i]\,\big)_{i \in [\![1,n]\!]}$ est un système complet d'événements.
  Par la formule des probabilités totales :

$$\begin{aligned}
\mathbb{P}\big(\,[R = A]\,\big) &= \sum_{i=1}^{n} \mathbb{P}\big(\,[J = i] \cap [R = A]\,\big) \\
&= \sum_{i=1}^{n} \mathbb{P}\big(\,[J = i] \cap [V \cup \{J\} = A]\,\big) && \textit{(par définition de } R\textit{)} \\
&= \sum_{i=1}^{n} \mathbb{P}\big(\,[J = i] \cap [V \cup \{i\} = A]\,\big) \\
&= \sum_{\substack{i\,=\,1 \\ i\,\in\,A}}^{n} \mathbb{P}\big(\,[J = i] \cap [V \cup \{i\} = A]\,\big) + \cancel{\sum_{\substack{i\,=\,1 \\ i\,\notin\,A}}^{n} \mathbb{P}\big(\,[J = i] \cap [V \cup \{i\} = A]\,\big)} && (*) \\
&= \sum_{i \in A} \mathbb{P}\big(\,[J = i] \cap [V \cup \{i\} = A]\,\big) && \textit{(car } A \subset [\![1,n]\!]\textit{)} \\
&= \sum_{i \in A} \mathbb{P}\big(\,[J = i]\,\big) \times \mathbb{P}_{[J=i]}\big(\,[V \cup \{i\} = A]\,\big) && \textit{(car pour tout } i \in [\![1,n]\!], \\
& && \mathbb{P}\big(\,[J = i]\,\big) \neq 0\textit{)}
\end{aligned}$$

- Pour conclure, il suffit alors de démontrer que pour tout $i \in A$:

$$\mathbb{P}_{[J=i]}\big(\,[V \cup \{i\} = A]\,\big) \;=\; \mathbb{P}_{[J=i]}\big(\,[V = A \setminus \{i\}]\,\big)$$

  Soit $i \in A$.

  × Si l'événement $[J = i]$ est réalisé, c'est qu'on a initialement tiré le numéro i (via un choix aléatoire dans $[\![1,n]\!]$).

  × Dans ce cas, l'événement $[V \cup \{i\} = A]$ est réalisé si et seulement si on obtient l'ensemble A en ajoutant i à la partie (qui ne contient pas i) obtenue par un tirage aléatoire simultané de $m - 1$ éléments. Cela est réalisé si et seulement si cette partie à $m - 1$ éléments coïncide avec l'ensemble A auquel on a retiré l'élément i.

$$\boxed{\text{On en conclut : } \forall i \in [\![1,n]\!],\ \mathbb{P}_{[J=i]}\big(\,[V \cup \{i\} = A]\,\big) \;=\; \mathbb{P}_{[J=i]}\big(\,[V = A \setminus \{i\}]\,\big).}$$

- Il reste enfin à justifier $(*)$.
  Pour ce faire, il suffit de remarquer que si $i \notin A$ alors $[V \cup \{i\} = A] = \varnothing$.

$$\boxed{\begin{array}{c} \text{On a donc bien, pour tout } A \in \mathcal{P}_m : \\ \mathbb{P}\big(\,[R = A]\,\big) \;=\; \displaystyle\sum_{i \in A} \mathbb{P}(\,[J = i]\,)\, \mathbb{P}_{[J=i]}\big(\,[V = A \setminus \{i\}]\,\big) \end{array}}$$

□

(ii) En déduire :

$$\mathbb{P}(\,[R = A]\,) \;=\; \frac{1}{\binom{n}{m}}\ \frac{\bar{x}_A}{\bar{x}}$$

Démonstration.
D'après la question précédente :

$$\begin{aligned}
\mathbb{P}\big(\,[R = A]\,\big) &= \sum_{i\in A} \mathbb{P}\big(\,[J = i]\,\big) \times \mathbb{P}_{[J=i]}\big(\,[V = A \setminus \{i\}]\,\big) \\
&= \sum_{i\in A} \mathbb{P}\big(\,[J = i]\,\big) \times \frac{1}{\binom{n-1}{m-1}} && \textit{(par définition de la loi conditionnelle de } V \textit{ relativement à l'événement } [J = i]\textit{)} \\
&= \frac{1}{\binom{n-1}{m-1}} \sum_{i\in A} \mathbb{P}\big(\,[J = i]\,\big) \\
&= \frac{1}{\binom{n-1}{m-1}} \ \sum_{i\in A} \ \frac{x_i}{\sum\limits_{r=1}^{n} x_r} && \textit{(par définition de la loi de } J\textit{)} \\
&= \frac{1}{\binom{n-1}{m-1}} \ \frac{1}{\sum\limits_{r=1}^{n} x_r} \ \sum_{i\in A} x_i \\
&= \frac{\color{red}{m}}{\color{blue}{n}\,\binom{n-1}{m-1}} \ \frac{\color{blue}{n}}{\sum\limits_{r=1}^{n} x_r} \ \frac{\sum\limits_{i\in A} x_i}{\color{red}{m}} \\
&= \frac{m}{n\,\binom{n-1}{m-1}} \ \frac{1}{\bar{x}_A} \ \bar{x}
\end{aligned}$$

Il suffit alors de remarquer :

$$n\binom{n-1}{m-1} \;=\; m\binom{n}{m} \quad \text{et ainsi} \quad \frac{1}{n\,\binom{n-1}{m-1}} \;=\; \frac{1}{m\,\binom{n}{m}} \quad \text{puis} \quad \frac{m}{n\,\binom{n-1}{m-1}} \;=\; \frac{1}{\binom{n}{m}}$$

$$\boxed{\text{On a bien, pour tout } A \in \mathcal{P}_m : \mathbb{P}(\,[R = A]\,) \;=\; \frac{1}{\binom{n}{m}}\ \frac{\bar{x}_A}{\bar{x}}.}$$

Commentaire

La relation sur les coefficients binomiaux de cette question est à considérer comme classique et il n'y a donc par lieu d'en faire la démonstration. On peut toutefois rappeler rapidement la manière de procéder. La première méthode est d'effectuer un simple calcul.

- Tout d'abord : $m\binom{n}{m} \;=\; m\,\dfrac{n!}{m!\,(n-m)!} \;=\; \dfrac{n!}{(m-1)!\,(n-m)!}$.
- De plus : $n\binom{n-1}{m-1} \;=\; n\,\dfrac{(n-1)!}{(m-1)!\,\big((n-\cancel{1})-(m-\cancel{1})\big)!} \;=\; \dfrac{n!}{(m-1)!\,(n-m)!}$.

Commentaire

Cette relation peut aussi se démontrer par dénombrement.
Pour ce faire, on considère un ensemble E à n éléments.
(on peut penser à une pièce qui contient n individus)
On souhaite alors construire une partie P à m éléments de cet ensemble contenant 1 élément distingué *(on peut penser à choisir dans la pièce un groupe de m individus dans lequel figure 1 représentant de ces individus)*.
Pour ce faire, on peut procéder de deux manières :

1) On choisit d'abord la partie à m éléments de E : $\binom{n}{m}$ possibilités.

On distingue ensuite 1 élément de cet ensemble P : $\binom{m}{1}$ possibilités.

(on choisit d'abord les m individus et on élit ensuite 1 représentants de ces individus)

Ainsi, il y a $\binom{n}{m}\binom{m}{1} = \binom{n}{m}\, m$ manières de construire P.

2) On choisit d'abord, dans E, l'élément à distinguer : $\binom{n}{1}$ possibilités.

On choisit ensuite $m-1$ éléments dans E, pour former P, en y ajoutant l'élément précédent : $\binom{n-1}{i-1}$ possibilités.

(on choisit d'abord 1 représentant puis on leur adjoint un groupe de $m-1$ individus)

Ainsi, il y a $\binom{n}{1}\binom{n-1}{m-1} = n\binom{n-1}{m-1}$ manières de construire P.

On retrouve ainsi le résultat souhaité. □

13. Une fois choisi le groupe de clients R (par la procédure de la question ***12.***), on définit : $\hat{\theta}_R = \dfrac{\bar{y}_R}{\bar{x}_R}$.

a) Démontrer :

$$\mathbb{E}\left(\hat{\theta}_R\right) = \frac{1}{\binom{n}{m}} \sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\bar{x}}$$

Démonstration.

- Remarquons tout d'abord, comme en question ***11.a)*** :

$$\hat{\theta}_R = \sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\bar{x}_A}\, \mathbb{1}_{[R=A]}$$

On peut notamment en déduire que la variable aléatoire $\hat{\theta}_R$ admet une espérance en tant que combinaison linéaire de variables aléatoires qui en admettent une.

- On a alors :

$$\begin{aligned}
\mathbb{E}(\hat{\theta}_R) &= \mathbb{E}\left(\sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\bar{x}_A}\, \mathbb{1}_{[R=A]}\right) \\
&= \sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\bar{x}_A}\, \mathbb{E}(\mathbb{1}_{[R=A]}) && \textit{(par linéarité de l'espérance)} \\
&= \sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\bar{x}_A}\, \mathbb{P}(\,[R=A]\,) && \textit{(car, pour tout } A\in\mathcal{P}_m, \ \mathbb{1}_{[R=A]} \hookrightarrow \mathcal{B}(\mathbb{P}([R=A]))\textit{)}
\end{aligned}$$

$$\boxed{\text{On en déduit, d'après la question précédente : } \mathbb{E}(\hat{\theta}_R) = \sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\cancel{\bar{x}_A}}\, \frac{1}{\binom{n}{m}}\, \frac{\cancel{\bar{x}_A}}{\bar{x}} = \frac{1}{\binom{n}{m}} \sum_{A\in\mathcal{P}_m} \frac{\bar{y}_A}{\bar{x}}.}$$

□

b) Conclure : $\mathbb{E}\left(\hat{\theta}_R\right) = \theta$. On a donc construit un estimateur non biaisé de θ.

Démonstration.

- On obtient alors :

$$\begin{array}{rcll}
\mathbb{E}(\hat{\theta}_R) & = & \dfrac{1}{\binom{n}{m}} \displaystyle\sum_{A \in \mathcal{P}_m} \dfrac{\bar{y}_A}{\bar{x}} & \\[2ex]
& = & \dfrac{\frac{1}{\binom{n}{m}} \displaystyle\sum_{A \in \mathcal{P}_m} \bar{y}_A}{\bar{x}} & \\[2ex]
& = & \dfrac{\mathbb{E}(Y)}{\bar{x}} & \textit{(en raisonnant comme en question } \textbf{\textit{11.a)(i)}}\textit{)} \\[2ex]
& = & \dfrac{\bar{y}}{\bar{x}} & \textit{(d'après la question } \textbf{\textit{11.a)(iv)}}\textit{)} \\[2ex]
& = & \dfrac{\cancel{\frac{1}{n}} \displaystyle\sum_{i=1}^{n} y_i}{\cancel{\frac{1}{n}} \displaystyle\sum_{i=1}^{n} x_i} \ = \ \theta &
\end{array}$$

$$\boxed{\text{On a bien : } \mathbb{E}(\hat{\theta}_R) = \theta.}$$

□