Concours Blanc 1 – Maths 2

ESSEC2, 2020

Corrigé

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 ($\Omega, \mathscr{A}, \mathbf{P}$) (ou $(\Omega, \mathscr{A}, \pi)$ dans la dernière question du sujet). Pour toute variable aléatoire X, on notera $\mathbf{E}(X)$ son espérance et $\mathbf{V}(X)$ sa variance lorsqu'elles existent.

Partie I. Biais par la taille, exemples discrets.

1. **On suppose que le nombre d'enfants dans une famille française est modélisé par une variable aléatoire X (lorsque l'on choisit une famille au hasard dans l'ensemble de toutes les familles françaises). 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. On suppose que X suit la loi binomiale de paramètres $n = 10$ et $p = 1/5$. On note $p_k = \mathbf{P}(X = k)$ pour $k \in \{0,1,\dots,10\}$.

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

   Pour $X \hookrightarrow \mathscr{B}\left(10, \frac{1}{5}\right)$, on a : $\forall\, k \in [\![0,10]\!]$; $\mathbf{P}(X = k) = \binom{10}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{10-k}$.

   ii. **Donner $\mathbf{E}(X)$ et $\mathbf{V}(X)$, et en déduire $\mathbf{E}\left(X^2\right)$.**

   D'après le cours, $\mathbf{E}(X) = 10 \times \frac{1}{5} = 2$; et $\mathbf{V}(X) = 10 \times \frac{1}{5} \times \frac{4}{5} = \frac{8}{5}$.

   Par Konig-Huygens, $\mathbf{E}(X^2) = \mathbf{V}(X) + \mathbf{E}(X)^2 = \frac{8}{5} + 4 = \frac{28}{5}$.

   (b) **Soit M_k le nombre de familles à k enfants dans la population.**

   Soit de plus $M = \sum_{k=0}^{10} M_k$ le nombre total de familles dans la population. On suppose que les proportions observées sur les familles des élèves de l'école s'identifient à la loi de probabilité de X : on a donc $p_k = \frac{M_k}{M}$. 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 $N = \sum_{k=0}^{10} N_k$ le nombre total d'enfants de la population.

   i. **Montrer que : $N_k = k p_k M$.**

   N_k s'obtient en sommant les enfants des familles à k enfants. Il y a M_k telles familles : on a donc $N_k = kM_k$.

   Or $p_k = \frac{M_k}{M}$: on a bien $N_k = k p_k M$.

   ii. **Montrer que : $\frac{N}{M} = \mathbf{E}(X)$.**

On a $N = \sum_{k=1}^{10} N_k = \sum_{k=1}^{10} kp_k M = M\sum_{k=1}^{10} kp_k$ et on reconnaît dans cette dernière somme l'espérance de X.
Donc finalement $N = M\mathbf{E}(X)$ ce qui donne le résultat attendu.

iii. **Montrer que la proportion des enfants provenant d'une famille à k enfants est : $p_k^* = \dfrac{kp_k}{2}$.**
La proportion recherchée est égale à $\dfrac{N_k}{N} = kp_k\dfrac{M}{N} = \dfrac{kp_k}{2}$.
On a bien $p_k^* = \dfrac{kp_k}{\mathbf{E}(X)} = \dfrac{kp_k}{2}$.

(c) **On suppose toujours que les proportions observées sur les familles des élèves de cette école s'identifient à celles de la population générale. On choisit un élève au hasard dans l'école, 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 k élément de $\{1,2,\dots,10\}$, justifier que : $\mathbf{P}(Y = k) = \dfrac{kp_k}{2}$.**
L'écolier interrogé fait partie d'une famille à k enfants avec une probabilité p_k^* (d'après les proportions de la question précédente).
On a donc : $\forall\, k \in [\![1,10]\!],\ \mathbf{P}(Y_k) = p_k^* = \dfrac{kp_k}{2}$.
(NB : pour $k = 0$ cette proba est évidemment nulle : il sera difficile de trouver un écolier issu d'une famille à 0 enfant.)

ii. **Montrer que : $\mathbf{E}(Y) = \dfrac{\mathbf{E}\left(X^2\right)}{\mathbf{E}(X)}$.**

$$\mathbf{E}(Y) = \sum_{k=1}^{10} k\mathbf{P}(Y = k) = \frac{1}{2}\sum_{k=1}^{10} k^2 p_k = \frac{1}{2}\sum_{k=0}^{10} k^2\mathbf{P}(X = k) = \frac{1}{2}\mathbf{E}(X^2) = \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$$

où le départ à $k = 0$ ne change pas la valeur de la somme (terme $k = 0$ nul) ; mais permet d'appliquer le théorème de transfert pour faire apparaître $\mathbf{E}(X^2)$. On se souvient enfin que $\mathbf{E}(X) = 2$.

iii. **En déduire la valeur de $\mathbf{E}(Y)$ et la comparer à $\mathbf{E}(X)$.**
Avec les valeurs trouvées plus haut : $\mathbf{E}(Y) = \dfrac{14}{5} > \dfrac{10}{5} = \mathbf{E}(X)$.

2. **Soit X une variable aléatoire à valeurs dans $\mathbb{N}$, non identiquement nulle et admettant une espérance non nulle. Pour tout entier naturel non nul i, on pose :**

$$q_i = \frac{i}{\mathbf{E}(X)}\mathbf{P}(X = i)$$

(a) **Montrer que la suite $\left(q_i\right)_{i\geqslant 1}$ définit une loi de probabilité.**

Les q_i ainsi définis sont positifs.
De plus :

$$\sum_{i=1}^{+\infty} q_i = \frac{1}{\mathbf{E}(X)}\sum_{i=1}^{+\infty} i\mathbf{P}(X = i) = \frac{1}{\mathbf{E}(X)}\sum_{i=0}^{+\infty} i\mathbf{P}(X = i) = \frac{1}{\mathbf{E}(X)}\mathbf{E}(X) = 1$$

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

On considère la variable aléatoire X^* dont la loi est donnée par les q_i, c'est-à-dire :

$$\forall\, i \in \mathbb{N}^*, \mathbf{P}\left(X^* = i\right) = \frac{i}{\mathbf{E}(X)}\mathbf{P}(X = i)$$

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

(b) **On suppose que X admet un moment d'ordre 2.**

i. **Montrer que: $\mathbf{E}\left(X^*\right) = \dfrac{\mathbf{E}\left(X^2\right)}{\mathbf{E}(X)}$.**

Pour avoir l'existence de $\mathbf{E}(X^*)$ on étudie la convergence (absolue (mais tout est positif ici) de la série de terme général $i\mathbf{P}(X^* = i) = \dfrac{1}{\mathbf{E}(X)} i^2 \mathbf{P}(X = i)$; cette dernière série converge bien car X admet un moment d'ordre 2.
On a donc l'existence de $\mathbf{E}(X^*)$, et

$$\mathbf{E}(X^*) = \sum_{i=1}^{+\infty} i\mathbf{P}(X^* = i) = \frac{1}{\mathbf{E}(X)} \sum_{i=1}^{+\infty} i^2 \mathbf{P}(X = i) = \frac{1}{\mathbf{E}(X)} \mathbf{E}(X^2)$$

ii. **En déduire que $\mathbf{E}(X^*) \geqslant \mathbf{E}(X)$.**
$\mathbf{E}(X^*) - \mathbf{E}(X) = \dfrac{1}{\mathbf{E}(X)}(\mathbf{E}(X^2) - \mathbf{E}(X)^2) = \dfrac{\mathbf{V}(X)}{\mathbf{E}(X)} \geqslant 0$ car une variance est positive.

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. **Expliciter la loi de X^*.**
On applique les formules : $X^*(\Omega) = \mathbb{N}^*$, et pour tout $i \in \mathbb{N}^*$:

$$\mathbf{P}(X^* = i) = \frac{i}{\mathbf{E}(X)} \mathbf{P}(X = i) = \frac{i}{\lambda} e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \frac{\lambda^{i-1}}{(i-1)!}$$

ii. **Vérifier que X^* suit la même loi que $X+1$.**
$X(\Omega) = \mathbb{N}$ donc $(X+1)(\Omega) = \mathbb{N}^*$; et pour tout $i \in \mathbb{N}^*$:

$$\mathbf{P}(X + 1 = i) = \mathbf{P}(X = i - 1) = e^{-\lambda} \frac{\lambda^{i-1}}{(i-1)!}$$

On voit donc que X^* et $X+1$ suivent la 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 $\forall k \in \mathbb{N}, \mathbf{P}(X = k+1) = \dfrac{\mathbf{E}(X)}{k+1} \mathbf{P}(X = k)$.**
Pour $k \in \mathbb{N}$, on a $\mathbf{P}(X = k) = \mathbf{P}(X + 1 = k + 1) = \mathbf{P}(X^* = k + 1) = \dfrac{k+1}{\mathbf{E}(X)} \mathbf{P}(X = k+1)$ et donc

$$\mathbf{P}(X = k+1) = \frac{\mathbf{E}(X)}{k+1} \mathbf{P}(X = k)$$

ii. **Montrer que $\forall k \in \mathbb{N}, \mathbf{P}(X = k) = \dfrac{(\mathbf{E}(X))^k}{k!} \mathbf{P}(X = 0)$.**
C'est naturellement une récurrence.
Pour $k = 0$ la propriété à démontrer s'écrit $\mathbf{P}(X = 0) = \dfrac{(\mathbf{E}(X))^0}{0!} \mathbf{P}(X = 0) = \dfrac{1}{1} \mathbf{P}(X = 0)$: c'est en effet vrai.

Soit maintenant $k \in \mathbb{N}$; on suppose que $\mathbf{P}(X = k) = \dfrac{(\mathbf{E}(X))^k}{k!} \mathbf{P}(X = 0)$. Alors d'après ce qui précède :

$$\mathbf{P}(X = k+1) = \frac{\mathbf{E}(X)}{k+1} \mathbf{P}(X = k) = \frac{\mathbf{E}(X)}{k+1} \frac{(\mathbf{E}(X))^k}{k!} \mathbf{P}(X = 0) = \frac{(\mathbf{E}(X))^{k+1}}{(k+1)!} \mathbf{P}(X = 0)$$

et on a bien la propriété au rang $k+1$.

Par récurrence, on a bien le résultat attendu.

iii. **En déduire la loi de X.**
Ça commence à ressembler à une loi de Poisson (et c'est bien normal car l'énoncé parle d'une «réciproque» à ce qui a été démontré en 3a). Notons $\mathbf{E}(X) = \lambda$. On a vu :

$$\forall k \in \mathbb{N},\ \mathbf{P}(X = k) = \frac{\lambda^k}{k!} \mathbf{P}(X = 0)$$

Il reste à déterminer $\mathbf{P}(X = 0)$... c'est un facteur multiplicatif qui se réglera si on utilise que la somme des probas vaut 1 !

$$\sum_{k=0}^{+\infty} \mathbf{P}(X = k) = \mathbf{P}(X = 0) \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = \mathbf{P}(X = 0)e^{\lambda} = 1$$

donc $\mathbf{P}(X = 0) = e^{-\lambda}$ et

$$\forall\, k \in \mathbb{N},\ \mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

X suit bien $\mathscr{P}(\lambda)$.

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 $\{1, 2, \dots, n\}$ telle que pour $1 \leqslant k \leqslant n$, $\mathbf{P}(X = k) > 0$. 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 de bus 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 $\{1, \dots, n\}$) et on suppose que pour tout entier k élément de $\{1, \dots, n\}$:**

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

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

   Dans cette somme sur j, seuls les termes $1 \leqslant j \leqslant k$ sont non nuls. On écrit donc en utilisant les probabilités données :

   $$\sum_{j=1}^{n} j\,\mathbf{P}_{(X=k)}(T = j) = \sum_{j=1}^{k} j\,\mathbf{P}_{(X=k)}(T = j) = \sum_{j=1}^{k} j\frac{1}{k} = \frac{1}{k}\sum_{j=1}^{k} j = \frac{1}{k}\frac{k(k+1)}{2} = \frac{k+1}{2}$$

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

   $$\sum_{k=1}^{n}\sum_{j=1}^{n} j\,\underbrace{\mathbf{P}(X = k)}_{\text{indép de } j}\,P_{(X=k)}(T = j) = \sum_{k=1}^{n} \mathbf{P}(X = k) \sum_{j=1}^{n} j\,P_{(X=k)}(T = j)$$

   et on injecte ce qu'on a trouvé à la question précédente :

   $$\sum_{k=1}^{n}\sum_{j=1}^{n} j\,\mathbf{P}(X = k)\,\mathbf{P}_{(X=k)}(T = j) = \sum_{k=1}^{n} \mathbf{P}(X = k)\frac{k+1}{2} = \frac{1}{2}\sum_{k=1}^{n}(k+1)\mathbf{P}(X = k) = \frac{\mathbf{E}(X+1)}{2}$$

   la dernière étape venant du théorème de transfert (pas de souci d'existence des espérances ici, les variables prennent un nombre fini de valeurs).

   iii. **En déduire que : $\mathbf{E}(T) = \dfrac{\mathbf{E}(X)+1}{2}$.**

   On a $\mathbf{E}(T) = \displaystyle\sum_{j=1}^{n} j\mathbf{P}(T = j)$, et avec les probas totales sur le SCE $\Big((X = k)\Big)_{1\leqslant k\leqslant n}$:

   $$\mathbf{P}(T = j) = \sum_{k=1}^{n} \mathbf{P}(X = k)\,\mathbf{P}_{(X=k)}(T = j)$$

   En reportant :

   $$\mathbf{E}(T) = \sum_{j=1}^{n} j\mathbf{P}(T = j) = \sum_{j=1}^{n} j \sum_{k=1}^{n} \mathbf{P}(X = k)\,\mathbf{P}_{(X=k)}(T = j) = \sum_{j=1}^{n}\sum_{k=1}^{n} j\mathbf{P}(X = k)\,\mathbf{P}_{(X=k)}(T = j)$$

   On peut intervertir les symboles $\sum_j$ et $\sum_k$ pour retomber sur la somme double de la question précédente !
   On en déduit immédiatement

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

(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 à considérer est ainsi X^*, suivant la loi de X biaisée par la taille. Le temps d'attente T' correspondant à cette situation vérifie donc en fait, pour $k \in \{1, \dots, n\}$:**

$$\mathbf{P}_{[X^*=k]}\left(T' = j\right) = \begin{cases} \frac{1}{k} & \text{si } j \in \{1, \dots, k\} \\ 0 & \text{si } j > k \end{cases}$$

i. **Justifier que: $\mathbf{E}\left(T'\right) = \dfrac{\mathbf{E}(X^*)+1}{2}$**

Il s'agit exactement des mêmes calculs qu'en 4a, en mettant des *.

ii. **En déduire que $\mathbf{E}\left(T'\right) \geqslant \mathbf{E}(T)$.**
On voit que

$$\mathbf{E}(T^*) \geqslant \mathbf{E}(T) \Leftrightarrow \frac{\mathbf{E}(X^*+1)}{2} \geqslant \frac{\mathbf{E}(X+1)}{2} \Leftrightarrow \frac{\mathbf{E}(X^*)+1}{2} \geqslant \frac{\mathbf{E}(X)+1}{2} \Leftrightarrow \mathbf{E}(X^*) \geqslant \mathbf{E}(X)$$

(avec la linéarité de l'espérance pour $\mathbf{E}(X+1) = \mathbf{E}(X)+1$). Cette dernière propriété a été prouvée en 2(b)ii : il convient de voir si on est dans les hypothèses qui sont :

- X à valeurs dans $\mathbb{N}$,
- non nulle,
- d'espérance non nulle,
- admettant un moment d'ordre 2.

C'est ici le cas : $X(\Omega) = [\![1, n]\!]$ donc elle est bien à valeurs entières, est non nulle, d'espérance non nulle car X ne prend que des valeurs $\geqslant 1$, et elle admet une variance car elle ne prend qu'un nombre fini de valeurs.

On peut donc appliquer 2(b)ii et on conclut.

(c) **On souhaite vérifier les résultats théoriques précédents à l'aide d'une simulation numérique en Python. On suppose ici que $X = Y+1$, où Y suit la loi binomiale $\mathscr{B}\left(10, \frac{1}{2}\right)$.**

i. **Calculer $\mathbf{E}(T)$ et $\mathbf{E}\left(T'\right)$ dans cette situation et vérifier que $\mathbf{E}\left(T'\right) - \mathbf{E}(T) \simeq 0,2$.**
$Y \hookrightarrow \mathscr{B}\left(10, \frac{1}{2}\right)$ donc $\mathbf{E}(Y) = 5$ et $\mathbf{E}(X) = 6$; donc $\mathbf{E}(T) = \frac{\mathbf{E}(X)+1}{2} = \frac{7}{2}$.
De plus $\mathbf{V}(Y) = \frac{5}{2}$; donc $\mathbf{V}(X) = \frac{5}{2}$ puis $\mathbf{E}(X^2) = \mathbf{E}(X)^2 + \mathbf{V}(X) = 36 + \frac{5}{2} = \frac{77}{2}$.
On en déduit que $\mathbf{E}(X^*) = \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)} = \frac{77}{12}$ puis $\mathbf{E}(T') = \frac{\mathbf{E}(X^*)+1}{2} = \frac{89}{24}$.
$\mathbf{E}(T') - \mathbf{E}(T) = \frac{5}{24} \simeq 0,208$.

ii. **On suppose que les bus opèrent entre 8 h 01 et 18 h 00 et que la personne souhaitant prendre un bus arrive à un instant aléatoire t_0 entre 8 h 01 et 18 h 00 .**
Justifier pourquoi la fonction Python suivante permet de simuler la variable aléatoire donnant le temps d'attente de la personne à l'arrêt de bus :

```python
def T():
    t0 = rd.randint(1,601)
    s = 0
    while s <= t0:
        X = rd.binomial(10,1/2)+1
        s = s+X
    return s-t0
```

Dans ce programme, l'instant t_0 est un entier de $[\![1, 600]\!]$ choisi au hasard qui représente la minute d'arrivée du client à l'arrêt de bus (dans 10h de service il y a 600 minutes !).
La variable `s` prend pour valeurs successives les temps de passage des bus : au début `s` vaut 0, puis on lui ajoute un tirage de X : c'est le temps d'arrivée du premier bus. Puis on lui ajoute un autre tirage de X qui donne l'intervalle entre le premier et le second bus : `s` vaut alors le temps d'arrivée du second bus. etc.
On boucle jusqu'à ce que `s` dépasse `t0` : alors `s` vaudra le temps d'arrivée du premier bus que peut prendre notre client ; et donc `s-t0` est bien son temps d'attente.

iii. **Écrire un programme utilisant la fonction précédente qui calcule et affiche la moyenne du temps d'attente de la personne lorsque l'on réalise 10000 fois l'expérience.**

Banalement on renvoie la moyenne de 10000 résultats obtenus :

```
print(np.mean([T() for k in range(10000)]))
```

iv. **Trois exécutions successives de ce programme renvoient les valeurs `3.6824`, `3.7169` et `3.6855`. Qui de T ou de T' semble répondre au mieux à la situation considérée ?**

$\mathbf{E}(T) = \frac{7}{2} = 3.5$; et on obtient des résultats autour de $3.7 = 3.5 + 0.2$ donc autour de $\mathbf{E}(T')$. Il semble bel et bien que T' donne la bonne modélisation ici.

(d) **Pourquoi l'hypothèse d'indépendance des intervalles de temps entre deux bus consécutifs du modèle décrit en introduction de la question 4 paraît trop simplificatrice par rapport à la réalité ?**

Si un intervalle entre 2 bus varie c'est parce qu'un bus est impacté par un certain événement : mais alors cela aura un effet sur deux intervalles de temps (entre notre bus et celui d'avant ; et entre notre bus et celui d'après) !

Partie II. Propriétés du biais par la taille dans le cas discret.

Dans cette partie, X est une variable aléatoire discrète à valeurs dans $\mathbb{N}$ admettant une espérance $\mathbf{E}(X) > 0$. On rappelle que l'on dit que X^* suit la loi de X biaisée par la taille si :

$$\forall i \in \mathbb{N}^*, \mathbf{P}\left(X^* = i\right) = \frac{i}{\mathbf{E}(X)}\mathbf{P}(X = i)$$

5. **Dans cette question, on se fixe $f : \mathbb{R}_+ \to \mathbb{R}$ et $g : \mathbb{R}_+ \to \mathbb{R}$ deux fonctions croissantes. On suppose de plus que les espérances $\mathbf{E}(f(X)), \mathbf{E}(g(X))$ et $\mathbf{E}(f(X)g(X))$ existent.**

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

- Si $x_1 \leqslant x_2$, par croissance de f et g on a : $f(x_1) \leqslant f(x_2)$ et $g(x_1) \leqslant g(x_2)$; donc

$$\underbrace{\left(f(x_1) - f(x_2)\right)}_{\leqslant 0}\underbrace{\left(g(x_1) - g(x_2)\right)}_{\leqslant 0} \geqslant 0$$

- Si $x_1 > x_2$, on a cette fois $f(x_1) \geqslant f(x_2)$ et $g(x_1) \geqslant g(x_2)$, donc

$$\underbrace{\left(f(x_1) - f(x_2)\right)}_{\geqslant 0}\underbrace{\left(g(x_1) - g(x_2)\right)}_{\geqslant 0} \geqslant 0$$

On a bien l'inégalité recherchée dans tous les cas.

(b) **Soient X_1 et X_2 deux variables aléatoires indépendantes de même loi que X. Montrer que : $\mathbf{E}\left(\left(f(X_1) - f(X_2)\right)\left(g(X_1) - g(X_2)\right)\right) = 2\mathbf{E}(f(X)g(X)) - 2\mathbf{E}(f(X))\mathbf{E}(g(X))$.**

On commence par développer :

$$\begin{aligned}\mathbf{E}\Big(\left(f(X_1) - f(X_2)\right)\left(g(X_1) - g(X_2)\right)\Big) &= \mathbf{E}(f(X_1)g(X_1) - f(X_2)g(X_1) - f(X_1)g(X_2) + f(X_2)g(X_2)) \\ &= \mathbf{E}(f(X_1)g(X_1)) - \mathbf{E}(f(X_2)g(X_1)) - \mathbf{E}(f(X_1)g(X_2)) + \mathbf{E}(f(X_2)g(X_2)) \\ &\text{(par linéarité de l'espérance)}\end{aligned}$$

Ici, $f(X)g(X)$ admet une espérance, X_1 et X_2 sont de même loi que X donc $f(X_1)g(X_1)$ et $f(X_2)g(X_2)$ admettent une espérance ;
$f(X_1)$ et $g(X_2)$ admettent une espérance, sont indépendantes par lemme des coalitions, donc $f(X_1)g(X_2)$

admet une espérance et de même $f(X_2)g(X_1)$.
De plus on a

$$\mathbf{E}(f(X_1)g(X_1)) = \mathbf{E}(f(X_2)g(X_2)) = \mathbf{E}(f(X)g(X))$$

et

$$\mathbf{E}(f(X_1)) = \mathbf{E}(f(X_2)) = \mathbf{E}(f(X))$$

et

$$\mathbf{E}(g(X_1)) = \mathbf{E}(g(X_2)) = \mathbf{E}(g(X))$$

En reprenant le calcul :

$$\begin{aligned}\mathbf{E}\Big(\big(f(X_1)-f(X_2)\big)\big(g(X_1)-g(X_2)\big)\Big) &= \mathbf{E}(f(X_1)g(X_1)) - \mathbf{E}(f(X_2)g(X_1)) - \mathbf{E}(f(X_1)g(X_2)) + \mathbf{E}(f(X_2)g(X_2))\\ &= \mathbf{E}(f(X)g(X)) - \mathbf{E}(f(X_2))\mathbf{E}(g(X_1)) - \mathbf{E}(f(X_1))\mathbf{E}(g(X_2)) + \mathbf{E}(f(X)g(X))\\ &= 2\mathbf{E}(f(X)g(X)) - 2\mathbf{E}(f(X))\mathbf{E}(g(X))\end{aligned}$$

ce qui permet de conclure.

(c) **En déduire que : $\mathbf{E}(f(X)g(X)) \geqslant \mathbf{E}(f(X))\mathbf{E}(g(X))$.**

D'après 5a, la variable $\big(f(X_1)-f(X_2)\big)\big(g(X_1)-g(X_2)\big)$ ne prend que des valeurs positives ; donc est d'espérance positive. On en déduit que

$$\mathbf{E}(\big(f(X)\,g(X)\big) - \mathbf{E}(\big(f(X)\big)\,\mathbf{E}(\big(g(X)\big) \geqslant 0$$

ce qui permet de conclure.

6. (a) **Dans cette question, f et g sont deux fonctions définies sur $\mathbb{R}_+$ telles que : $\forall x \in \mathbb{R}_+,\ |f(x)| \leqslant g(x)$. Montrer que si $g(X)$ admet une espérance, alors $f(X)$ admet une espérance.**

C'est un problème de séries. Notons $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$; les x_k sont tous positifs par hypothèse.
Par théorème de transfert, on sait que $\sum_k g(x_k)\mathbf{P}(X = x_k)$ converge absolument, et on veut montrer que $\sum_k f(x_k)\mathbf{P}(X = x_k)$ converge absolument.
Mais, comme les x_k sont positifs, l'hypothèse de l'énoncé donne $0 \leqslant |f(x_k)| \leqslant g(x_k)$, et donc

$$\forall\, k \in \mathbb{N},\ 0 \leqslant |f(x_k)|\mathbf{P}(X = x_k) \leqslant g(x_k)\mathbf{P}(X = x_k)$$

($\mathbf{P}(X = x_k)$ étant bien sûr positif).
Par comparaison de SATP on en déduit la convergence absolue de $\sum_k f(x_k)\mathbf{P}(X = x_k)$ et donc l'existence de $\mathbf{E}(f(X))$.

(b) **Dans cette question, h est une fonction bornée définie sur $\mathbb{R}_+$.**

i. **Montrer que $h(X)$ admet une espérance.**
On a l'existence d'un $M \geqslant 0$ tel que $\forall\, x \geqslant 0,\ |h(x)| \leqslant M$. On remarque que la v.a. constante égale à M admet une espérance. En prenant pour g la fonction constante égale à M, on applique la question précédente et on en déduit que $h(X)$ admet une espérance.

ii. **Montrer que $Xh(X)$ admet une espérance et que : $\mathbf{E}(h(X^*)) = \dfrac{1}{\mathbf{E}(X)}\mathbf{E}(Xh(X))$.**

7. **Dans cette question, on suppose qu'il existe un entier $m \geqslant 1$ tel que $\mathbf{E}\left(X^{m+1}\right)$ existe.**

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

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

ii. **En déduire que $\mathbf{E}(X^p)$ existe.**
Par hypothèse $\mathbf{E}(X^{m+1})$ existe, donc $\mathbf{E}(1+X^{m+1}) = 1+\mathbf{E}(X^{m+1})$ aussi. X étant discrète positive, on déduit de $0 \leqslant x^p \leqslant 1+x^{m+1}$ et de la question 6a que $\mathbf{E}(X^p)$ existe.

(b) **Montrer que: $\mathbf{E}\left(X^{m+1}\right) \geqslant \mathbf{E}(X)\mathbf{E}(X^m)$.**
On applique la question 5c avec les fonctions $f : x \mapsto x$ et $g : x \mapsto x^m$ qui sont bien croissantes sur $\mathbb{R}_+$: $\mathbf{E}(X \times X^m) \geqslant \mathbf{E}(X)\,\mathbf{E}(X^m)$, ce qui donne bien :

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

(c) **Montrer que: $\mathbf{E}\left((X^*)^m\right) \geqslant \mathbf{E}(X^m)$.**
On reprend la loi de X^* et on applique le théorème de transfert : sous réserve de convergence absolue (mais tout est positif) :

$$\mathbf{E}((X^*)^m) = \sum_{k=1}^{+\infty} k^m \mathbf{P}(X^* = k) = \sum_{k=1}^{+\infty} k^m \frac{k}{\mathbf{E}(X)} \mathbf{P}(X = k) = \frac{1}{\mathbf{E}(X)} \sum_{k=1}^{+\infty} k^{m+1}\mathbf{P}(X = k) = \frac{\mathbf{E}(X^{m+1})}{\mathbf{E}(X)}$$

et on a donc bien convergence, car $\mathbf{E}(X^{m+1})$ existe ; donc

$$\mathbf{E}((X^*)^m) = \frac{\mathbf{E}(X^{m+1})}{\mathbf{E}(X)} \geqslant \frac{\mathbf{E}(X^m)\mathbf{E}(X)}{\mathbf{E}(X)} = \mathbf{E}(X^m)$$

8. **Pour tout t réel positif, on définit la fonction g_t sur $\mathbb{R}_+$ par :**

$$\forall x \in \mathbb{R}_+, g_t(x) = \begin{cases} 0 & \text{si } 0 \leqslant x \leqslant t \\ 1 & \text{si } x > t \end{cases}$$

(a) **Montrer que la fonction $x \mapsto g_t(x)$ est croissante sur $\mathbb{R}_+$.**
Le graphe de g_t est un créneau : fonction nulle sur $[0,t]$ et égale à 1 sur $]t,+\infty[$.
On voit donc que c'est croissant... mais ça ne constitue pas une preuve !
Faute de continuité et donc de pouvoir dériver, on revient à la définition : soient $(x, y) \in \mathbb{R}_+^2$ avec $x \leqslant y$.

- Si $x \leqslant y \leqslant t$ alors $g_t(x) = g_t(y) = 0$ et donc $g_t(x) \leqslant g_t(y)$;
- Si $x \leqslant t < y$ alors $g_t(x) = 0$ et $g_t(y) = 1$ donc $g_t(x) \leqslant g_t(y)$;
- et si $t < x \leqslant y$ alors $g_t(x) = g_t(y) = 1$ et donc $g_t(x) \leqslant g_t(y)$.

Dans tous les cas on a bien $g_t(x) \leqslant g_t(y)$ ce qui établit la croissance de g_t.

(b) **Montrer que pour tout t réel positif, $\mathbf{E}\left(Xg_t(X)\right)$ est bien définie et : $\mathbf{E}\left(Xg_t(X)\right) \geqslant \mathbf{E}(X)\mathbf{P}(X > t)$.**
On applique 5c avec $f(x) = x$ et $g(x) = g_t(x)$. f et g sont bien croissantes ; donc il faut étudier l'existence des espérances $\mathbf{E}\big(X\,g_t(X)\big)$, $\mathbf{E}(X)$ (celle-ci existe par hypothèse) et $\mathbf{E}\big(g_t(X)\big)$.
On procède par majoration : $g_t(x) = 0$ ou 1 donc X étant à valeurs positives on a $0 \leqslant Xg_t(X) \leqslant X$ et donc par la question 6a, $Xg_t(X)$ admet une espérance ; et avec $0 \leqslant g_t(X) \leqslant 1$ et le fait que la variable constante égale à 1 admet une espérance, on a aussi l'existence de l'espérance de $g_t(X)$.
Avec 5c on peut conclure que $\mathbf{E}\big(X\,g_t(X)\big) \geqslant \mathbf{E}(X)\mathbf{E}\big(g_t(X)\big)$.

Montrons maintenant que $\mathbf{E}\big(g_t(X)\big) = \mathbf{P}(X > t)$. Par transfert :

$$\mathbf{E}\big(X\,g_t(X)\big) = \sum_{k=0}^{+\infty} g_t(k)\mathbf{P}(X = k)$$

et $g_t(k)$ valant 0 si $k \leqslant t$ et 1 si $k > t$, on obtient

$$\mathbf{E}\big(X\,g_t(X)\big) = \sum_{k>t} \mathbf{P}(X = k) = \mathbf{P}(X > t)$$

Finalement on a bien démontré :

$$\mathbf{E}\big(X\,g_t(X)\big) \geqslant \mathbf{E}(X)\,\mathbf{P}(X > t)$$

(c) **En déduire que:** $\forall t \in \mathbb{R}, \mathbf{P}(X^* > t) \geqslant \mathbf{P}(X > t)$.
On dit que X^* **domine stochastiquement** X.
On utilise cette fois la question 6(b)ii , avec $h = g_t$ définie sur $\mathbb{R}_+$ et bornée par 1 (elle vaut 0 ou 1...).

On a donc

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

Comme dans la question précédente on montre que $\mathbf{E}\big(g_t(X^*)\big) = \mathbf{P}(X^* > t)$; et avec $\mathbf{E}\big(X\,g_t(X)\big) \geqslant \mathbf{E}(X)\,\mathbf{P}(X > t)$ on obtient

$$\mathbf{P}(X^* > t) = \mathbf{E}\big(g_t(X^*)\big) = \frac{1}{\mathbf{E}(X)}\,\mathbf{E}\big(X\,g_t(X)\big) \geqslant \frac{1}{\mathbf{E}(X)}\mathbf{E}(X)\,\mathbf{P}(X > t) = \mathbf{P}(X > t)$$

ce qui fournit bien l'inégalité demandée.

Partie III. Généralisation au cas de variables aléatoires quelconques.

Soit X **est une variable aléatoire positive quelconque, non nécessairement discrète.**
On suppose que X **admet une espérance** $\mathbf{E}(X) > 0$**.**
Sur le modèle de la question 6(b)ii, on dit que X^* **suit la loi de** X **biaisée par la taille si :**

$$\text{Pour toute fonction } h : \mathbb{R}_+ \to \mathbb{R} \text{ bornée, } \mathbf{E}\left(h\left(X^*\right)\right) = \frac{1}{\mathbf{E}(X)}\mathbf{E}(Xh(X))$$

On admet que cette propriété caractérise une unique loi de probabilité.
Soit n **un entier naturel non nul et soient** $X_1, X_2, \ldots, X_n$ **des variables aléatoires positives quelconques, indépendantes, non nécessairement de même loi.**
On suppose que pour tout $i \in \{1, 2, \ldots, n\}, X_i$ **admet une espérance strictement positive et on note** $\mu_i = \mathbf{E}(X_i)$**.**
On pose de plus :

$$\mu = \sum_{i=1}^{n} \mu_i \text{ et } S_n = \sum_{i=1}^{n} X_i$$

9. **Calculer** $\mathbf{E}(S_n)$**.**

Les X_i admettent une espérance donc S_n aussi, et par linéarité on a immédiatement :

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

10. **Pour** A **un événement, on note** $\mathbb{1}_A$ **la variable aléatoire définie par** $\mathbb{1}_A(\omega) = 1$ **si** $\omega \in A$**, et** $\mathbb{1}_A(\omega) = 0$ **sinon.**

(a) **Montrer que** $\mathbb{1}_A$ **suit la loi de Bernoulli de paramètre** $\mathbf{P}(A)$**.**
$\mathbb{1}_A$ vaut 0 ou 1 ; et $(\mathbb{1}_A = 1) = A$ (égalité entre événements) d'où $\mathbf{P}(\mathbb{1}_A = 1) = \mathbf{P}(A)$; ceci donne bien le résultat.

(b) **Montrer que si** $(A_1, A_2, \ldots, A_n)$ **est un système complet d'événements, alors** $\sum_{i=1}^{n} \mathbb{1}_{A_i}$ **est la variable aléatoire constante égale à 1.**
Soit $\omega \in \Omega$ quelconque : par propriété d'un SCE on sait que ω appartient à un et un seul des A_i. Si on note A_{i_0} cet événement, on a

$$\sum_{i=1}^{n} \mathbb{1}_{A_i}(\omega) = \mathbb{1}_{A_{i_0}}(\omega) = 1$$

On a bien montré que $\sum_{i=1}^{n} \mathbb{1}_{A_i} = 1$.

11. **On considère des variables aléatoires indépendantes $X_1^*, X_2^*, \dots, X_n^*$, indépendantes des X_i, telles que pour tout $i \in \{1,2,\dots,n\}, X_i^*$ suit la loi de X_i biaisée par la taille.**
Soit J une variable aléatoire indépendante de $X_1, X_1^*, X_2, X_2^*, \dots, X_n, X_n^*$, de loi donnée par :

$$\forall i \in \{1,2,\dots,n\}, \mathbf{P}(J = i) = \frac{\mu_i}{\mu}$$

On considère la variable aléatoire $X_J = \sum_{i=1}^{n} X_i \mathbb{1}_{[J=i]}$ et on définit $T_n = S_n - X_J + X_J^*$.

Autrement dit, on choisit un indice aléatoire J et, dans la somme $\sum_{i=1}^{n} X_i$, on remplace X_J par X_J^*.

Soit enfin $h : \mathbb{R}_+ \to \mathbb{R}$ une fonction bornée.

(a) i. **Montrer que : $h(T_n) = \sum_{i=1}^{n} h\left(S_n - X_i + X_i^*\right) \mathbb{1}_{[J=i]}$.**

On considère le SCE associé à J : $\left((J = i)\right)_{1 \leqslant i \leqslant n}$; avec la question 10 on peut écrire $\sum_{i=1}^{n} \mathbb{1}_{(J=i)} = 1$, et donc

$$h(T_n) = \sum_{i=1}^{n} h(T_n) \mathbb{1}_{(J=i)}$$

Soit maintenant $\omega \in \Omega$:

- si $\omega \in (J = i)$, $(J = i)$ est réalisé ; alors $T_n = S_n - X_i + X_i^*$ et $\mathbb{1}_{(J=i)} = 1$, donc $h(T_n)\mathbb{1}_{(J=i)} = h(S_n - X_i + X_i^*)\mathbb{1}_{(J=i)}$.
- si $\omega \notin (J = i)$, $(J = i)$ n'est pas réalisé, donc $\mathbb{1}_{(J=i)} = 0$ et donc $h(T_n)\mathbb{1}_{(J=i)} = h(S_n - X_i + X_i^*)\mathbb{1}_{(J=i)} = 0$.

Dans tous les cas :

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

et on a donc bien montré :

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

ii. **En déduire que : $\mathbf{E}(h(T_n)) = \sum_{i=1}^{n} \mathbf{P}(J = i)\mathbf{E}\left(h\left(S_n - X_i + X_i^*\right)\right)$.**

On considère $h(T_n) = \sum_{i=1}^{n} h(S_n - X_i + X_i^*)\mathbb{1}_{(J=i)}$, et on va essayer de prendre l'espérance.

La fonction h est bornée sur $\mathbb{R}_+$; on déduit de la question 6(b)i l'existence de l'espérance de $h(T_n)$ et des $h(S_n - X_i + X_i^*)$.
D'autre part l'indicatrice $\mathbb{1}_{(J=i)}$ admet aussi une espérance égale à $\mathbf{P}(J = i)$.
De plus, J est indépendante des X_i, X_i^* donc par lemme des coalitions $\mathbb{1}_{(J=i)}$ est indépendante des $h(S_n - X_i + X_i^*)$.
On peut donc écrire :

$$\begin{aligned}
\mathbf{E}(h(T_n)) &= \sum_{i=1}^{n} \mathbf{E}\left(h(S_n - X_i + X_i^*)\mathbb{1}_{(J=i)}\right) \\
&= \sum_{i=1}^{n} \mathbf{E}\left(h(S_n - X_i + X_i^*)\right)\mathbf{E}\left(\mathbb{1}_{(J=i)}\right) \\
&= \sum_{i=1}^{n} \mathbf{P}(J = i)\mathbf{E}\left(h(S_n - X_i + X_i^*)\right)
\end{aligned}$$

(b) **Pour $i \in \{1,2,\dots,n\}$, démontrer que : $\forall s \in \mathbb{R}_+, \mathbf{E}\left(h\left(s + X_i^*\right)\right) = \frac{1}{\mu_i}\mathbf{E}(X_i h(s + X_i))$.**

On admettra qu'on en déduit l'égalité : $\mathbf{E}\left(h\left(S_n - X_i + X_i^*\right)\right) = \frac{1}{\mu_i}\mathbf{E}(X_i h(S_n))$.

On cherche à utiliser la caractérisation de la loi biaisée par la taille (équation de 6(b)ii). La fonction $h_s : x \mapsto h(s + x)$ est bornée sur $\mathbb{R}_+$ car h l'est ; donc

$$\mathbf{E}\left(h_s(X_i^*)\right) = \frac{1}{\mathbf{E}(X_i)} \mathbf{E}\left(X_i h_s(X_i)\right)$$

ce qui s'écrit encore :

$$\mathbf{E}\left(h(s + X_i^*)\right) = \frac{1}{\mu_i} \mathbf{E}\left(X_i h(s + X_i)\right)$$

(c) **En déduire que : $\mathbf{E}(h(\mathrm{T}_n)) = \dfrac{\mathbf{E}(\mathrm{S}_n h(\mathrm{S}_n))}{\mathbf{E}(\mathrm{S}_n)}$.**

On compile plusieurs résultats précédents :

$$\begin{aligned}
\mathbf{E}\big(h(\mathrm{T}_n)\big) &= \sum_{i=1}^{n} \mathbf{P}(\mathrm{J}=i)\mathbf{E}\big(h(\mathrm{S}_n - \mathrm{X}_i + \mathrm{X}_i^*)\big) \\
&= \sum_{i=1}^{n} \frac{\mu_i}{\mu}\frac{1}{\mu_i}\mathbf{E}\big(\mathrm{X}_i\, h(\mathrm{S}_n)\big) \qquad \text{(loi de J, + résultat admis)} \\
&= \frac{1}{\mu}\sum_{i=1}^{n} \mathbf{E}\big(\mathrm{X}_i\, h(\mathrm{S}_n)\big) \\
&= \frac{1}{\mu}\mathbf{E}\left(\sum_{i=1}^{n} \mathrm{X}_i\, h(\mathrm{S}_n)\right) \qquad \text{(linéarité de l'espérance)} \\
\mathbf{E}\big(h(\mathrm{T}_n)\big) &= \frac{\mathbf{E}(\mathrm{S}_n\, h(\mathrm{S}_n))}{\mathbf{E}(\mathrm{S}_n)} \qquad (\mathbf{E}(\mathrm{S}_n) = \mu)
\end{aligned}$$

(d) **Conclure que T_n suit la loi de S_n biaisée par la taille.**
Ce qu'on a fait est valable pour toute fonction $h : \mathbb{R}_+ \to \mathbb{R}$ bornée : ce fait résulte donc de la définition donnée par l'énoncé !

Partie IV. Une application en statistiques.

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-ème client a payé x_i euros ($x_i > 0$), mais a en réalité consommé une quantité d'électricité correspondant à y_i euros $\big(y_i > 0\big)$. La compagnie sait combien ses clients ont payé, et elle souhaite estimer le rapport :

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

pour déterminer à quel point elle a mal facturé ses clients.
En pratique elle ne peut pas sonder tout le monde : elle décide donc de choisir m clients parmi les n pour effectuer ses mesures.
On considère ainsi dans cette partie l'univers $\Omega = \mathscr{P}_m$ constitué des parties à m éléments de $\{1, 2, \ldots, n\}$.
Pour tout $\mathrm{A} \in \mathscr{P}_m$, on note $\{\mathrm{A}\}$ l'événement : « On choisit la partie A de $\{1, 2, \ldots, n\}$ » . Si l'on suppose que pour sonder les clients on choisit une partie de $\{1, 2, \ldots, n\}$ de manière uniforme dans $\mathscr{P}_m$, on a alors :

$$\forall \mathrm{A} \in \mathscr{P}_m, \mathbf{P}\big(\{\mathrm{A}\}\big) = \frac{1}{\binom{n}{m}}$$

Pour $\mathrm{A} \in \mathscr{P}_m$, on notera :

$$\overline{x_\mathrm{A}} = \frac{1}{m}\sum_{i \in \mathrm{A}} x_i \text{ et } \overline{y_\mathrm{A}} = \frac{1}{m}\sum_{i \in \mathrm{A}} y_i$$

On note également :

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ et } \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

On définit enfin deux variables aléatoires X et Y sur $\Omega = \mathscr{P}_m$ de la manière suivante :

$$\forall \mathrm{A} \in \mathscr{P}_m, \mathrm{X}(\mathrm{A}) = \overline{x_\mathrm{A}} \quad \text{et} \quad \mathrm{Y}(\mathrm{A}) = \overline{y_\mathrm{A}}$$

La compagnie d'électricité décide d'utiliser $\theta_1 = \dfrac{\mathrm{Y}}{\mathrm{X}}$ comme estimateur de θ.
On admet que si l'on munit un univers fini Ω de la probabilité P, alors l'espérance d'une variable aléatoire T définie sur Ω est donnée par:

$$\mathbf{E}(\mathrm{T}) = \sum_{\omega\in\Omega} \mathrm{T}(\omega)\mathbf{P}\left(\{\omega\}\right)$$

12. (a) **Que représentent les nombres $\overline{x}$ et $\overline{y}$?**

    $\overline{x}$ est la facturation moyenne des clients ; et $\overline{y}$ leur consommation (réelle) moyenne.

    (b) **Que représentent les variables aléatoires X et Y ?**

    À une issue (donc une partie de $[\![1,n]\!]$ à m éléments) notée A, la variable X associe la facturation moyenne des clients de l'ensemble A ; et Y associe la consommation moyenne des clients de A.

13. (a) **Montrer que : $\mathbf{E}(\mathrm{X}) = \dfrac{1}{\binom{n}{m}} \displaystyle\sum_{\mathrm{A}\in\mathscr{P}_m} \overline{x_\mathrm{A}}$.**

    On prend la formule admise pour l'espérance : $\mathbf{E}(\mathrm{X}) = \displaystyle\sum_{\mathrm{A}\in\mathscr{P}_m} \mathrm{X}(\mathrm{A})\mathbf{P}\left(\{\mathrm{A}\}\right) = \sum_{\mathrm{A}\in\mathscr{P}_m} \overline{x_\mathrm{A}}\frac{1}{\binom{n}{m}}$

    (b) **Soit $i \in \{1,2,\ldots,n\}$ un entier fixé.**
    Combien y a-t-il de parties $\mathrm{A} \in \mathscr{P}_m$ telles que $i \in \mathrm{A}$?

    Pour construire une telle partie, il faut :

    - choisir i
    - choisir les $m-1$ autres éléments de A parmi les $n-1$ restants (on ne peut pas choisir deux fois i !). Ceci donne $\binom{n-1}{m-1}$ possibilités.

    Le nombre de parties recherché est donc $\binom{n-1}{m-1}$.

    (c) **En déduire que $\displaystyle\sum_{\mathrm{A}\in\mathscr{P}_m}\sum_{i\in\mathrm{A}} x_i = \binom{n-1}{m-1}\sum_{i=1}^{n} x_i$.**

    Soit $i_0 \in [\![1,n]\!]$ quelconque. Dans la double somme $\displaystyle\sum_{\mathrm{A}\in\mathscr{P}_m}\sum_{i\in\mathrm{A}} x_i$, x_{i_0} apparaît autant de fois que de parties de $\mathscr{P}_m$ qui le contiennent ; donc $\binom{n-1}{m-1}$ fois d'après la question précédente.

    Ceci valant pour n'importe quel i_0, on a :

$$\sum_{\mathrm{A}\in\mathscr{P}_m}\sum_{i\in\mathrm{A}} x_i = \binom{n-1}{m-1}x_1 + \binom{n-1}{m-1}x_2 + \cdots + \binom{n-1}{m-1}x_n = \binom{n-1}{m-1}\sum_{i=1}^{n} x_i$$

    (d) **Conclure que $\mathbf{E}(\mathrm{X}) = \overline{x}$. On admettra que de même on a $\mathbf{E}(\mathrm{Y}) = \overline{y}$.**

    On reprend les résultats précédents et les définitions de $\overline{x}_\mathrm{A}$ et $\overline{x}$.

$$\begin{aligned}
\mathbf{E}(\mathrm{X}) &= \frac{1}{\binom{n}{m}} \sum_{\mathrm{A}\in\mathscr{P}_m} \overline{x}_\mathrm{A} \\
&= \frac{1}{\binom{n}{m}} \sum_{\mathrm{A}\in\mathscr{P}_m} \frac{1}{m} \sum_{i\in\mathrm{A}} x_i \\
&= \frac{1}{m\binom{n}{m}} \sum_{\mathrm{A}\in\mathscr{P}_m}\sum_{i\in\mathrm{A}} x_i \\
&= \frac{1}{m\binom{n}{m}} \binom{n-1}{m-1}\sum_{i=1}^{n} x_i \\
&= \frac{1}{m\binom{n}{m}} \binom{n-1}{m-1} n\overline{x}
\end{aligned}$$

    Avec des factorielles :

$$\frac{1}{m\binom{n}{m}}\binom{n-1}{m-1}n = \frac{1}{m}\,\frac{m!(n-m)!}{n!} \times \frac{(n-1)!}{(m-1)!(n-m)!} \times n = 1$$

    et on conclut donc $\mathbf{E}(\mathrm{X}) = \overline{x}$.

(e) **Exprimer** θ **en fonction de** $\mathbf{E}(X)$ **et** $\mathbf{E}(Y)$**.**

On a $\theta = \dfrac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \dfrac{n\overline{y}}{n\overline{x}} = \dfrac{\overline{y}}{\overline{x}} = \dfrac{\mathbf{E}(Y)}{\mathbf{E}(X)}$ d'après les calculs qui précèdent.

14. **Soient** W **et** Z **deux variables aléatoires strictement positives admettant un moment d'ordre 2. On note, pour tout** $t \in \mathbb{R} : Q(t) = \mathbf{E}\left((W + tZ)^2\right)$**.**

(a) **Montrer que** Q **est un polynôme du second degré et déterminer le signe de** Q **sur** $\mathbb{R}$**.**

W et Z admettent un moment d'ordre 2 donc les espérances de W^2, Z^2 et WZ existent.
Par linéarité de l'espérance, $Q(t)$ existe et $Q(t) = \mathbf{E}(W^2) + 2t\mathbf{E}(WZ) + t^2\mathbf{E}(Z^2)$.
Q est donc bien un trinôme.
De plus $(W + tZ)^2 \geqslant 0$ donc par positivité de l'espérance : $\forall\, t \geqslant 0,\ Q(t) \geqslant 0$.

(b) **En considérant le discriminant de** Q**, en déduire que** $(\mathbf{E}(WZ))^2 \leqslant \mathbf{E}\left(W^2\right)\mathbf{E}\left(Z^2\right)$**.**

Q est un trinôme de signe constant donc de discriminant négatif.
En reprenant l'expression : $\Delta = (2\mathbf{E}(WZ))^2 - 4\mathbf{E}(W^2)\mathbf{E}(Z^2) = 4(\mathbf{E}(WZ)^2 - \mathbf{E}(W^2)\mathbf{E}(Z^2))$.
$\Delta \leqslant 0$ donne l'inégalité recherchée.

(c) **Montrer que l'inégalité de la question précédente est une égalité si et seulement si il existe un réel** $\alpha > 0$ **tel que** $W = \alpha Z$ **presque sûrement.**

Avec les quantités de la question précédente : l'inégalité est une égalité ssi $\Delta = 0$; donc ssi Q admet une racine t_0. Pour ce t_0 on a donc $\mathbf{E}\left((W + t_0Z)^2\right) = 0$.
La variable $(W+t_0Z)^2$ est positive d'espérance nulle, donc presque sûrement nulle. Avec probabilité 1 on a donc $W + t_0Z = 0$ ou encore $W = -t_0Z$.
Enfin, la stricte positivité de W et Z montre que $\alpha = -t_0$ est forcément > 0.

NB : L'inégalité $(\mathbf{E}(WZ))^2 \leqslant \mathbf{E}\left(W^2\right)\mathbf{E}\left(Z^2\right)$ est appelée Inégalité de Cauchy-Schwarz, elle apparaît dans différents contextes en mathématiques (c'est notamment celle qui assure qu'un coefficient de corrélation linéaire est dans $[-1, 1]$). La cas d'égalité est lui aussi usuel.

15. (a) **Montrer que:** $\mathbf{E}\left(\dfrac{1}{X}\right) \geqslant \dfrac{1}{\mathbf{E}(X)}$**.**

Tous les x_i étant > 0, X est à valeurs strictement positives. On applique l'inégalité de la question 14b aux variables $\sqrt{X}$ et $\dfrac{1}{\sqrt{X}}$; ces variables admettent des moments de tous ordres car elles ne prennent qu'un nombre fini de valeurs.
On obtient :

$$\mathbf{E}\left(\sqrt{X}\,\frac{1}{\sqrt{X}}\right) \leqslant \sqrt{\mathbf{E}(X)}\sqrt{\mathbf{E}\left(\frac{1}{X}\right)}$$

donc $\mathbf{E}(X)\mathbf{E}\left(\dfrac{1}{X}\right) \geqslant 1$; ce qui s'écrit encore $\mathbf{E}\left(\dfrac{1}{X}\right) \geqslant \dfrac{1}{\mathbf{E}(X)}$ (avec $\mathbf{E}(X) > 0$)

(b) **Montrer qu'il y a égalité si et seulement si** X **est une variable aléatoire presque sûrement constante égale à** $\overline{x}$**.**

D'après 14c, il y a égalité dans l'inégalité précédente si et seulement s'il existe $\alpha > 0$ tel que $\sqrt{X} = \alpha\dfrac{1}{\sqrt{X}}$ avec proba 1, ce qui donne $X = \alpha$ avec proba 1.
Dans le cas d'une variable presque sûrement constante on sait qu'on a alors $\mathbf{E}(X)$ égale à cette constante ; et comme $\mathbf{E}(X) = \overline{x}$, on conclut que $\mathbf{E}\left(\dfrac{1}{X}\right) = \dfrac{1}{\mathbf{E}(X)}$ si et seulement si X est la variable aléatoire

presque sûrement constante égale à $\overline{x}$.

(c) **Conclure que $\mathbf{E}\left(\frac{1}{X}\right) = \frac{1}{\mathbf{E}(X)}$ si et seulement si $x_i = \overline{x}$ pour tout $i \in \{1, 2, \ldots, n\}$.**

Rappelons qu'une valeur de X est obtenue en prenant une partie A à m éléments au hasard dans $[\![1, n]\!]$ et en calculant la moyenne des x_i associés.
Supposons qu'il existe $i \neq j$ tels que $x_i \neq x_j$. Soit A_1 est une partie contenant i mais pas j, et A_2 la partie obtenue à partir de A_1 en remplaçant i par j.
Par exemple en supposant $i < j$, on peut prendre $A_1 = [\![1, i]\!]$ et $A_2 = [\![1, i-1]\!] \cup \{j\}$.
On voit alors que la valeur de X associée à A_1 sera différente de celle associée à A_2, car on aura remplacé x_i par $x_j \neq x_i$. Dans ce cas la variable X n'est pas constante p.s. (les deux événements mentionnés sont de probabilité égale à $\frac{1}{\binom{n}{m}} > 0$).
Ainsi, si X est p.s. constante alors tous les x_i sont égaux ; et $\overline{x}$ étant leur moyenne, ils sont tous égaux à $\overline{x}$.

16. **Si l'on suppose que X et Y sont indépendantes, montrer que $\mathbf{E}(\theta_1) \geqslant \theta$, avec égalité si seulement si $x_i = \overline{x}$ pour tout $i \in \{1, 2, \ldots, n\}$.**

On a $\mathbf{E}(\theta_1) = \mathbf{E}\left(\frac{Y}{X}\right) = \mathbf{E}\left(Y \times \frac{1}{X}\right)$; si X et Y sont indépendantes alors Y et $\frac{1}{X}$ le sont aussi par lemme des coalitions.
Alors

$$\mathbf{E}(\theta_1) = \mathbf{E}\left(Y \times \frac{1}{X}\right) = \mathbf{E}(Y) \times \mathbf{E}\left(\frac{1}{X}\right) \geqslant \mathbf{E}(Y) \times \left(\frac{1}{\mathbf{E}(X)}\right) = \frac{\mathbf{E}(Y)}{\mathbf{E}(X)} = \theta$$

On a égalité dans ce qui précède si et seulement si $\mathbf{E}(Y) = 0$ (exclu) ou $\mathbf{E}\left(\frac{1}{X}\right) = \frac{1}{\mathbf{E}(X)}$ ce qui, on l'a vu, correspond à $x_i = \overline{x}$ pour tout i.

Ainsi $\mathbf{E}(\theta_1)$ n'est pas forcément égal à θ : on dit que θ_1 est un estimateur biaisé de θ.
Ce problème peut être résolu en choisissant la partie A contenant les m clients étudiés non de manière uniforme comme dans les questions 13 à 16, mais de manière biaisée par la taille. Par analogie avec la construction de T_n dans la question 11, on commence par choisir un indice aléatoire J à valeurs dans $\{1, 2, \ldots, n\}$, dont la loi est donnée par :

$$\forall i \in \{1, 2, \ldots, n\}, \mathbf{P}(J = i) = \frac{x_i}{\sum_{r=1}^{n} x_r} \qquad (**)$$

Ensuite, sachant $[J = i]$, on choisit un groupe V de $m-1$ clients parmi les $n-1$ clients différents de i, de manière uniforme:

$$\mathbf{P}_{[J=i]}\left(\{V\}\right) = \frac{1}{\binom{n-1}{m-1}}$$

La partie choisie pour effectuer le sondage est alors: $A = V \cup \{i\}$.

17. **On souhaite programmer en Python un tirage aléatoire de la variable J, connaissant les valeurs des x_i.**

(a) **On rappelle qu'en Python, la commande `rd.random()` renvoie un tirage d'une variable X équirépartie dans $[0, 1]$: en particulier, pour tout $p \in [0, 1]$, $\mathbf{P}(X \leqslant p) = p$.**
Montrer que si $0 \leqslant x \leqslant y \leqslant 1$, $\mathbf{P}(X \in]x, y]) = y - x$.

Pour $0 \leqslant x \leqslant y \leqslant 1$ on peut écrire $(X \leqslant y) = (X \leqslant x) \cup (x < X \leqslant y)$ (union disjointe) ; d'où

$$\mathbf{P}(X \leqslant y) = \mathbf{P}(X \leqslant x) + \mathbf{P}(x < X \leqslant y)$$

Or $\mathbf{P}(X \leqslant y) = y$ et $\mathbf{P}(X \leqslant x) = x$ d'où

$$\mathbf{P}(X \in]x, y]) = y - x$$

(b) **On note, pour $k \in [\![0, n]\!]$: $r_0 = 0$, et $r_k = \frac{\sum_{r=1}^{k} x_i}{\sum_{r=1}^{n} x_r}$ si $k \geqslant 1$. On a en particulier $r_n = 1$.**
Un tirage de la variable X décrite en $17a$ étant effectué, on note j l'unique entier de $[\![1, n]\!]$ tel que

$r_{j-1} < \mathrm{X} \leqslant r_j$ **(on ne demande pas de justifier l'existence et l'unicité de** j**).**
Montrer que l'indice j **défini ainsi est un tirage de la variable aléatoire J dont la loi est donnée par** $(**)$**.**

Cet indice j est à valeurs dans $[\![1,n]\!]$.
Une valeur de X étant tirée, l'indice défini par cette procédure vaudra j ssi $r_{j-1} < \mathrm{X} \leqslant r_j$; événement dont on vient de voir qu'il est vrai avec probabilité $r_j - r_{j-1} = \dfrac{\sum_{r=1}^{j} x_i}{\sum_{r=1}^{n} x_r} - \dfrac{\sum_{r=1}^{j-1} x_i}{\sum_{r=1}^{n} x_r} = \dfrac{x_j}{\sum_{r=1}^{n} x_r} = \mathbf{P}(\mathrm{J} = j)$.
L'indice j ainsi défini suit bien la même loi que la variable aléatoire J.

(c) **Coder une fonction Python** `def tirageJ(X)` **qui prend en argument un** `np.array` $\mathrm{X} = [x_1, \ldots, x_n]$ **et renvoie un tirage de la variable aléatoire J.** On propose :

```
def tirageJ(L):
    x = rd.random()
    s = 0
    j = 0
    while s<x:
        s = s + L[j]/np.sum(L)
        j = j+1
    return j
```

Dans ce code on tire un x =`rd.random()` ; la variable `s` vaut successivement $r_1, r_1 + r_2, \ldots$. jusqu'à $r_1 + \ldots + r_n = 1$ et on renvoie le premier indice j tel que $r_j \geqslant x$, ce qui est la caractérisation demandée.

(d) **On suppose codée une fonction** `partie_aleatoire(L,m)` **qui prend en argument une liste L de taille** p**, et un entier** $m \leqslant p$**, et renvoie une liste de** m **élements de L tirée aléatoirement de manière uniforme.**
Programmer, à l'aide d'un code qui appellera les deux fonctions précédentes, une fonction `tirage_taille(n,m,X)` **qui permet de tirer une partie** $\mathrm{A} \subset [\![1,n]\!]$ **de cardinal** `m`**, de manière biaisée par la taille,** `X` **étant la liste des** x_i**.**
Rappel : **si** `L` **est une liste de longueur** n**, et** $0 \leqslant a \leqslant b \leqslant n$**,** `L[a:b]` **est la liste** `[L[a],L[a+1],...,L[b-1]]`**. De plus, si** `L` **et** `M` **sont deux listes, la commande** `L+M` **renvoie la concaténation de ces 2 listes.**
Si on relit l'énoncé, il faut dans un premier temps choisir un indice J par la fonction précédente ; puis compléter la partie par $m-1$ éléments choisis aléatoirement de manière uniforme dans « ce qui reste » .
On peut alors proposer :

```
def tirage_taille(n,X,m):
    j = tirageJ(X)
    L = list(range(1,n+1))
    L1=L[0:j-1]+L[j:n]
    return [j]+partie_aleatoire(L1,m-1)
```

Ici `L=[1,2,3,...,n]` et donc d'après le découpage en tranches exposé, `L[0:j-1]=[1,2,...,j-1]` et `L[j,n]=[j+1,...,n]` de sorte que `L[0:j-1]+L[j,n]` est `L` privée de l'élément `j`.
On tire donc les `m-1` éléments restants dans cette liste ; et on renvoie `j` et les `m-1` restants dans un même objet.

18. **Dans cette question, on détermine pour** $\mathrm{A} \in \mathscr{P}_m$**, la probabilité** p_{A} **que** A **soit choisie avec le protocole précédent.**

(a) **Montrer que** $p_{\mathrm{A}} = \sum_{i \in \mathrm{A}} \mathbf{P}(\mathrm{J} = i) \mathbf{P}_{[\mathrm{J}=i]}\left(\{\mathrm{A} \setminus \{i\}\}\right)$**.**

Si on note $\mathrm{A} = (i_1, \ldots, i_m)$, l'événement «choisir la partie A» est l'union disjointe des $\big((\mathrm{J} = i_k) \cap (\mathrm{V} = \mathrm{A} \setminus \{i_k\})\big)$.
Donc

$$\begin{aligned} p_{\mathrm{A}} &= \sum_{k=1}^{m} \big((\mathrm{J}=i_k)\cap(\mathrm{V}=\mathrm{A}\setminus\{i_k\}\big) \\ &= \sum_{i\in\mathrm{A}} \mathrm{P}\big((\mathrm{J}=i)\cap(\mathrm{V}=\mathrm{A}\setminus\{i\}\big) \\ &= \sum_{i\in\mathrm{A}} \mathbf{P}(\mathrm{J}=i)\mathrm{P}_{(\mathrm{J}=i)}\big(\mathrm{V}=\mathrm{A}\setminus\{i\}\big) \end{aligned}$$

(b) **En déduire que** $p_{\mathrm{A}} = \dfrac{1}{\binom{n}{m}}\dfrac{\overline{x_{\mathrm{A}}}}{\overline{x}}$**.**

Avec $\mathbf{P}(\mathrm{J}=i) = \dfrac{x_i}{\sum_{r=1}^{n} x_r}$ et $\mathrm{P}_{(\mathrm{J}=i)}(\mathrm{V}=\mathrm{A}\setminus\{i\}) = \dfrac{1}{\binom{n-1}{m-1}}$ on trouve

$$p_{\mathrm{A}} = \sum_{i\in\mathrm{A}} \frac{x_i}{\sum_{r=1}^{n} x_r}\frac{1}{\binom{n-1}{m-1}} = \frac{1}{\binom{n-1}{m-1}}\frac{1}{n\overline{x}}\underbrace{\sum_{i\in\mathrm{A}} x_i}_{=m\overline{x}_{\mathrm{A}}} = \frac{\overline{x}_{\mathrm{A}}}{\overline{x}}\frac{m}{n}\frac{1}{\binom{n-1}{m-1}}$$

et avec le dernier calcul de la question 13d on voit que $\dfrac{m}{n}\dfrac{1}{\binom{n-1}{m-1}} = \dfrac{1}{\binom{n}{m}}$ ce qui conclut.

19. **Dans cette dernière question, on munit à présent l'univers $\Omega = \mathscr{P}_m$ d'une nouvelle probabilité π définie par :**

$$\forall \mathrm{A}\in\mathscr{P}_m, \pi\left(\{\mathrm{A}\}\right) = \frac{1}{\binom{n}{m}}\frac{\overline{x_{\mathrm{A}}}}{\overline{x}}$$

Avec cette nouvelle distribution de probabilité, on reprend la variable aléatoire θ_1 telle que :

$$\forall \mathrm{A}\in\mathscr{P}_m, \theta_1(\mathrm{A}) = \frac{\mathrm{Y}(\mathrm{A})}{\mathrm{X}(\mathrm{A})} = \frac{\overline{y_{\mathrm{A}}}}{\overline{x_{\mathrm{A}}}}$$

(a) **Montrer que** $\mathbf{E}(\theta_1) = \sum_{\mathrm{A}\in\mathscr{P}_m}\pi\left(\{\mathrm{A}\}\right)\theta_1(\mathrm{A}) = \dfrac{1}{\binom{n}{m}}\sum_{\mathrm{A}\in\mathscr{P}_m}\dfrac{\overline{y_{\mathrm{A}}}}{\overline{x}}$**.**

On a, en appliquant la définition de l'espérance pour la proba π et avec ce qui précède :

$$\begin{aligned} \mathbf{E}(\theta_1) = \sum_{\mathrm{A}\in\mathscr{P}_m} \pi(\{\mathrm{A}\})\theta_1(\mathrm{A}) &= \sum_{\mathrm{A}\in\mathscr{P}_m} \frac{1}{\binom{n}{m}}\frac{\overline{x}_{\mathrm{A}}}{\overline{x}}\frac{\overline{y}_{\mathrm{A}}}{\overline{x}_{\mathrm{A}}} \\ &= \frac{1}{\binom{n}{m}}\sum_{\mathrm{A}\in\mathscr{P}_m}\frac{\overline{y}_{\mathrm{A}}}{\overline{x}} \end{aligned}$$

(b) **Conclure que** $\mathbf{E}(\theta_1) = \theta$**.**
En choisissant l'échantillon A à étudier de manière biaisée par la taille, θ_1 est cette fois un estimateur non biaisé de θ.
En continuant le calcul précédent :

$$\begin{aligned} \mathbf{E}(\theta_1) &= \frac{1}{\binom{n}{m}}\sum_{\mathrm{A}\in\mathscr{P}_m}\frac{\overline{y}_{\mathrm{A}}}{\overline{x}} \\ &= \frac{1}{\overline{x}}\frac{1}{\binom{n}{m}}\sum_{\mathrm{A}\in\mathscr{P}_m}\frac{\sum_{i\in\mathrm{A}} y_i}{m} \\ &= \frac{1}{\overline{x}}\frac{1}{\binom{n}{m}}\frac{1}{m}\left(\sum_{\mathrm{A}\in\mathscr{P}_m}\sum_{i\in\mathrm{A}} y_i\right) \\ &= \frac{1}{\overline{x}}\frac{1}{\binom{n}{m}}\frac{1}{m}\binom{n-1}{m-1}\sum_{i=1}^{n} y_i \qquad \text{d'après 13c avec des } y_i \text{ au lieu des } x_i \\ &= \frac{1}{\overline{x}}\frac{1}{\binom{n}{m}}\frac{1}{m}\binom{n-1}{m-1} n\overline{y} \end{aligned}$$

Avec l'identité vue en question 13d, on a $\frac{1}{\binom{n}{m}} \frac{1}{m} \binom{n-1}{m-1} n = 1$, d'où finalement :

$$\mathbf{E}(\theta_1) = \frac{\overline{y}}{\overline{x}} = \theta.$$