Devoir Surveillé 02

Le samedi 14 octobre 2023
durée 3h00

Documents écrits, électroniques, calculatrices et téléphones portables interdits
La plus grande attention sera apportée à la qualité de la rédaction, syntaxe et orthographe comprise.
Numérotez les copies et questions et soulignez ou encadrez les résultats.
Vous pouvez utiliser les résultats des questions en amont d'une question donnée en la référençant clairement
Les calculs, même longs et fastidieux doivent figurer sur la copie

Si au cours de l'épreuve, vous repérez ce qui vous semble être une erreur d'énoncé, vous le signalerez sur votre copie et poursuivrez la composition.

Exercice I

Dans cet exercice, on suppose donné un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$ et X, Y, Z des variables aléatoires à valeurs réelles, chacune prenant un nombre fini de valeurs. On note $\mathbb{E}(X)$ l'espérance d'une variable aléatoire réelle X et $\mathbb{V}(X)$ sa variance.

1. Rappeler la démonstration de l'identité de KOENIG–HUYGHENS :

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2$$

2. On définit la covariance de deux variables aléatoires X et Y comme étant :

$$\mathbb{C}ov(X,Y) = \mathbb{E}(X.Y) - \mathbb{E}(X).\mathbb{E}(Y)$$

2.a. Démontrer que :

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\,\mathbb{C}ov(X,Y)$$

puis que

$$\mathbb{V}(X+Y+Z) = \mathbb{V}(X) + \mathbb{V}(Y) + \mathbb{V}(Z) + 2\,\mathbb{C}ov(X,Y) + 2\,\mathbb{C}ov(Y,Z) + 2\,\mathbb{C}ov(X,Z)$$

2.b. Montrer que si X et Y sont indépendantes alors $\mathbb{C}ov(X,Y) = 0$.

2.c. Donner un exemple de deux variables aléatoires X et Y *non indépendantes* telles que $\mathbb{C}ov(X,Y) = 0$. On pourra choisir pour X une v.a. pouvant prendre trois valeurs (entières) et pour Y une fonction de X.

3.a. Démontrer que $\mathbb{C}ov(X,Y)^2 \leq \mathbb{V}(X).\mathbb{V}(Y)$. On pourra étudier la fonction d'une variable réelle f définie par

$$\forall t \in \mathbb{R},\, f(t) = \mathbb{V}(X + t.Y)$$

3.b. On suppose $\mathbb{V}(Y) \neq 0$. A quelle condition sur X et Y existe-t-il $t_0 \in \mathbb{R}$ tel que $f(t_0) = 0$?

3.c. A quelle condition sur X et Y a-t-on l'égalité :

$$\mathbb{C}ov(X,Y)^2 = \mathbb{V}(X).\mathbb{V}(Y)$$

4. Etant données X, Y et Z des variables aléatoires centrées (*i.e.* d'espérances nulles, $\mathbb{V}(X) = \mathbb{E}(X^2)$,...), on définit la fonction f sur $\mathbb{R}^2$ par la formule :

$$\forall (a,b) \in \mathbb{R}^2,\, f(a,b) = \mathbb{E}\left((Z - a.X - b.Y)^2\right)$$

4.a. Montrer que

$$\forall (a,b) \in \mathbb{R},\, f(a,b) = \mathbb{V}(Z) + a^2\mathbb{V}(X) + b^2\mathbb{V}(Y) - 2a.\,\mathbb{C}ov(X,Z) - 2b\,\mathbb{C}ov(Y,Z) + 2ab\,\mathbb{C}ov(X,Y)$$

4.b. On pose $\Delta = \mathbb{V}(X).\mathbb{V}(Y) - \mathbb{C}ov(X,Y)^2$ et on suppose que $\Delta \neq 0$. Déterminer le(s) point(s) critique(s) éventuel(s) de f.

Exercice II

Toutes les variables aléatoires sont supposées définies sur un même espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$.

Les réponses numériques qui sont des fractions rationnelles seront données sous forme réduite.

Une puce se déplace sur un axe gradué. À l'instant 0, la puce se trouve sur le point d'abscisse 0.

À partir de l'instant $n = 0$, la puce effectue à chaque instant, un saut vers la droite selon le protocole suivant :

— elle effectue un saut d'une unité vers la droite avec la probabilité $\frac{1}{2}$;

— elle effectue un saut de deux unités vers la droite avec la probabilité $\frac{1}{4}$;

— elle effectue un saut de trois unités vers la droite avec la probabilité $\frac{1}{4}$.

Les différents sauts sont supposés indépendants.

Pour tout entier $n \geq 1$, on définit les variables aléatoires suivantes :

— S_n est égale au nombre d'unités sautées lors du n-ième saut ;

— X_n est égale au nombre de sauts d'une unité effectués lors des n premiers sauts ;

— Y_n est égale au nombre de sauts de deux unités effectués lors des n premiers sauts ;

— Z_n est égale au nombre de sauts de trois unités effectués lors des n premiers sauts ;

— A_n est égale à l'abscisse du point occupé par la puce à l'issue de son n-ième saut.

1. Donner la loi de la variable aléatoire A_1. Calculer $\mathbb{E}(A_1)$ et $\mathbb{V}(A_1)$.

2.a. Vérifier que $A_2 = S_1 + S_2$ et calculer la loi du couple (S_1, S_2) puis de A_2. On présentera les résultats sous forme de tableaux.

2.b. Calculer $\mathbb{E}(A_2)$.

3. Compléter (en la recopiant sur la copie) la fonction suivante pour qu'elle simule (et retourne) les `N` (`N` vaut 100 par défaut) premiers déplacements de la puce.

```
def S(N=100):
    s = [0]*(N+1)
    for k in range(N):
        u = numpy.random.rand()
        if u <= ..... :
            s[k] = 1
        elif u <= ..... :
            s[k] = 2
        else :
            s[k] = 3
    return s
```

4. Pour $n \in \mathbb{N}^*$, reconnaître les lois de X_n, Y_n et Z_n. Montrer que $X_n + Y_n + Z_n = n$ et justifier que $X_n + Y_n$, suit la loi binomiale $\mathscr{B}\left(n, \frac{3}{4}\right)$.

5. Pour $n \in \mathbb{N}^*$, en utilisant les valeurs de $\mathbb{V}(X_n)$, $V(Y_n)$ et $\mathbb{V}(X_n + Y_n)$, déterminer la covariance $\mathbb{C}ov(X_n, Y_n)$.

Les v.a. X_n et Y_n sont-elles indépendantes ?

6.a. Exprimer A_n, en fonction de X_n, Y_n et Z_n et montrer que $\mathbb{E}(A_n) = \frac{7}{4}.n$.

6.b. Exprimer A_n, en fonction de X_n, et Y_n et calculer $\mathbb{V}(A_n)$. **6.c.** Calculer $\mathbb{C}ov(A_n, X_n)$.

7. On complète le programme Python de la question 3 en y ajoutant les lignes suivantes :

```
Sauts = S(N=200)
A = [0]*(len(Sauts))
for k in range(1,len(A)):
    A[k] = A[k-1] + Sauts[k]
abscisses = [k for k in range(len(A))]
matplotlib.pyplot.plot(abscisses,A)
```

Quelle sortie graphique obtient-on ?

En supposant que chaque `A[k]` est proche de l'espérance théorique $\mathbb{E}(A_k)$, quelle est l'allure de la courbe tracée (Faire un graphique explicatif.) ?

Exercice III

Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$. Si Y est une variable aléatoire *réelle* admettant une espérance, on note cette espérance $\mathbb{E}(Y)$.

Partie A

Préliminaires : Fonction logarithme en base 2 et propriété de convexité.

A.1. La notation ln désigne la fonction logarithme « népérien ». La fonction logarithme en base 2, notée $\log_2$, est définie sur $]0,+\infty[$ par la formule :

$$\forall x \in]0,+\infty[,\ \log_2(x) = \frac{\ln x}{\ln 2}$$

A.1.a. Montrer que

$$\forall (x,y) \in (]0,+\infty[)^2,\ \log_2(x.y) = \log_2(x) + \log_2(y).$$

A.1.b. Montrer que

$$\forall \alpha \in \mathbb{R},\ \log_2(2^\alpha) = \alpha.$$

A.1.c. Quel est le signe de la fonction $\log_2$ sur l'intervalle $]0,1[$?

A.2. Fonctions convexes sur un intervalle. Soit I un intervalle de $\mathbb{R}$ non vide et non réduit à un point. Une fonction $\phi : I \to \mathbb{R}$, de classe $\mathscr{C}^2$ sur I sera dite *convexe* sur I si

$$\forall x \in I, \phi''(x) \geq 0$$

A.2.a. Montrer que la fonction ϕ définie par $\forall t \in \mathbb{R},\ \phi(t) = 2^t$ est convexe sur $I = \mathbb{R}$.

A.2.b. Montrer que la fonction $\phi = -\log_2$ est convexe sur $I =]0,+\infty[$.

A.2.c. On suppose que $\phi : I \to \mathbb{R}$, de classe $\mathscr{C}^2$ sur l'intervalle non trivial I y est convexe. Montrer que :

$$\forall (x,y) \in I^2, \forall t \in [0,1],\ \phi(t.x + (1-t).y) \leq t.\phi(x) + (1-t).\phi(y)$$

Indication:On définira et étudiera (en construisant tableau de signes/variations de g'' et g') la fonction g définie, après avoir fixé $x, y \in I$, par

$$\forall t \in [0,1],\ g(t) = t.\phi(x) + (1-t).\phi(y) - \phi(t.x + (1-t).y)$$

On pourra aussi remarquer, en le justifiant, que g' s'annule au moins une fois sur $]0,1[$.

⋆⋆⋆

On admet l'énoncé suivant, se déduisant du précédent par une récurrence élaborée :

Si $\phi : I \to \mathbb{R}$, de classe $\mathscr{C}^2$ sur l'intervalle non trivial I, y est convexe alors

$$\forall n \in \mathbb{N}^*, \forall (x_1,\dots,x_n) \in I^n, \forall (t_1,\dots,t_n) \in ([0,1])^n, \sum_{i=1}^n t_i = 1 \Rightarrow \phi\left(\sum_{i=1}^n t_i.x_i\right) \leq \sum_{i=1}^n t_i.\phi(x_i) \qquad \text{(Cvx)}$$

Partie B

Entropie d'une variable aléatoire réelle discrète.

Soit $A \subset \mathbb{R}$ un ensemble fini de nombres réels non vide. On dit que X est une variable aléatoire dont la loi est à *support* A, si X est à valeurs dans A et si pour tout $x \in A$, $\mathbb{P}(X = x) > 0$.

B.1. Soit X une variable aléatoire de loi à support $\{0, 1, 2, \ldots, n\}$ où n est un entier naturel. On appelle *entropie probabiliste* de X le réel

$$H(X) = -\sum_{k=0}^{n} \mathbb{P}(X = k) \log_2 \mathbb{P}(X = k). \tag{+}$$

B.1.a. On définit la fonction $g : \{0, \ldots, n\} \to \mathbb{R}$ en posant $g(k) = \log_2 \mathbb{P}(X = k)$ pour k élément de $\{0, \ldots, n\}$. Montrer que $H(X) = -\mathbb{E}(g(X))$.

B.1.b. Montrer que $H(X) \geq 0$.

B.1.c. Soit p un réel tel que $0 < p < 1$. On suppose dans cette question que X suit la loi de BERNOULLI $\mathscr{B}(p)$.

B.1.c.i. Calculer $H(X)$ en fonction de p. On note ψ la fonction qui, à p, associe $H(X)$.

B.1.c.ii. Montrer que $-\psi$ est convexe sur $]0, 1[$.

B.1.c.iii. Déterminer la valeur p_0 où ψ est maximale.

B.1.d. On suppose dans cette question que la loi de X est à support $\{0, 1, 2, 3\}$ avec les probabilités

$$\mathbb{P}(X = 0) = \frac{1}{2}; \mathbb{P}(X = 1) = \frac{1}{4}; \mathbb{P}(X = 2) = \mathbb{P}(X = 3) = \frac{1}{8}.$$

Calculer $H(X)$.

B.2. On souhaite écrire une fonction `Entropie` en Python/Numpy pour calculer l'entropie d'une variable aléatoire X dont le support de la loi est de la forme $A = \{0, \ldots, n\}$ où n est un entier naturel. On suppose que le vecteur `P` passé en argument est tel que pour tout k de A, `P[k]` $= \mathbb{P}(X = k)$. Compléter le corps de la fonction ci-dessous d'argument `P` qui renvoie l'entropie de X, c'est-à-dire $-\sum_{k=0}^{n} \mathbb{P}(X = k) \log_2 \mathbb{P}(X = k)$.

```
def Entropie(P):
    """
    Retourne l'entropie de X v.a. à support {0,..,n} dont la loi est décrite
    par le vecteur P de longueur n+1
    """
```

Si nécessaire, on pourra utiliser l'instruction `len(P)` qui donne le nombre d'éléments de `P` et la fonction `np.log`, qui, une fois le module `numpy` importé sous l'alias `np`, calcule le logarithme *népérien* de son argument.

⋆⋆⋆

On souhaite maintenant démontrer quelques inégalités concernant l'entropie.

B.3. On commence par une inégalité générale, appelée Inégalité de JENSEN.

Soit $n \geq 1$. Soit X une variable aléatoire de loi à support $\{x_0, x_1, \ldots, x_n\}$ où les x_i sont des nombres distincts de $\mathbb{R}$. On pose $\mathbb{P}(X = x_i) = p_i$.

Montrer que pour tout $0 \leq i \leq n$, on a $p_i < 1$.

On suppose que I est un intervalle de $\mathbb{R}$ contenant les nombres x_i et $\phi : I \to \mathbb{R}$ est une fonction convexe. Justifier que $\mathbb{E}(X) \in I$ et que

$$\phi\left(\mathbb{E}(X)\right) \leq \mathbb{E}(\phi(X))$$

B.4. Soit X une variable aléatoire de loi à support $\{0, \dots, n\}$. On pose, pour k tel que $0 \leq k \leq n$,

$$p_k = \mathbb{P}(X = k).$$

B.4.a. Montrer que

$$\sum_{k=0}^{n} p_k \log_2 \frac{1}{(n+1)p_k} \leq \log_2 \left(\sum_{k=0}^{n} p_k \frac{1}{(n+1)p_k} \right) = 0$$

B.4.b. Montrer que

$$\sum_{k=0}^{n} p_k \log_2[(n+1)p_k] = \log_2(n+1) - H(X)$$

B.4.c. Montrer que $H(X) \leq \log_2(n+1)$.
B.4.d. On suppose que X suit la loi uniforme sur $\{0, \dots, n\}$. Calculer $H(X)$.
B.5. Soient X et Y deux variables aléatoires de même loi à support $\{0, \dots, n\}$. On suppose en outre X et Y indépendantes.
B.5.a. Montrer que $\mathbb{P}(X = Y) = \sum_{k=0}^{n} \mathbb{P}(X = k)^2$.
B.5.b. On pose $v(k) = \mathbb{P}(X = k)$ pour tout k élément de $\{0, \dots, n\}$. Montrer que

$$2^{\mathbb{E}(\log_2 v(X))} \leq \mathbb{E}(2^{\log_2 v(X)}) = \mathbb{E}(v(X)).$$

B.5.c. En déduire que $2^{-H(X)} \leq \mathbb{P}(X = Y)$.
B.5.d. Donner un exemple de loi où l'inégalité précédente est une égalité.

PYTHON
AGRO-VETO
2020

Listes

`[]`	Créer une liste vide
`[a]*n`	Créer une liste avec n fois l'élément `a`
`L.append(a)`	Ajoute l'élément `a` à la fin de la liste `L`
`L1 + L2`	Concatène les deux listes `L1` et `L2`
`len(L)`	Renvoie le nombre d'éléments de la liste `L`
`L.pop(k)`	Renvoie l'élément d'indice k de `L` et l'enlève de `L`
`L.remove(a)`	Enlève une fois la valeur `a` de la liste `L`
`max(L)`	Renvoie le plus grand élément de la liste `L`
`min(L)`	Renvoie le plus petit élément de la liste `L`
`sum(L)`	Renvoie la somme de tous les éléments de la liste `L`

Numpy

`import numpy as np`

`np.array()`	Transforme une liste en matrice `numpy`
`np.linspace(a,b,n)`	Crée une matrice ligne de n valeurs uniformément réparties entre a et b (inclus)
`np.zeros([n,m])`	Crée la matrice nulle de taille $n \times m$
`np.eye(n)`	Crée la matrice identité de taille n
`np.diag(L)`	Crée la matrice diagonale dont les termes diagonaux sont les éléments de la liste `L`
`np.transpose(M)`	Renvoie la transposée de M
`np.dot(M,P)`	Renvoie le produit matriciel MP
`np.sum(M)`	Renvoie la somme de tous les éléments de M
`np.prod(M)`	Renvoie le produit de tous les éléments de M
`np.max(M)`	Renvoie le plus grand élément de M
`np.min(M)`	Renvoie le plus petit élément de M
`np.shape(M)`	Renvoie dans un couple le format de la matrice M
`np.size(M)`	Renvoie le nombre d'éléments de M

Numpy.linalg

`import numpy.linalg as la`

`la.inv(M)`	Renvoie l'inverse de la matrice M si elle est inversible
`la.eigvals(M)`	Renvoie la liste des valeurs propres de M
`la.eig(M)`	Renvoie un couple `L,P` où `L` est la liste des valeurs propres de M et `P` la matrice de passage associée
`la.matrix_rank(M)`	Renvoie le rang de M

Random

`import numpy.random as rd`

`rd.rand()`	Simule une réalisation d'une variable $X \hookrightarrow \mathcal{U}([0,1])$
`rd.randint(a,b)`	Simule une réalisation d'une variable $X \hookrightarrow \mathcal{U}([\![a,b[\![)$
`rd.gauss(0,1)`	Simule une réalisation d'une variable $X \hookrightarrow \mathcal{N}(0,1)$
`rd.choice(L)`	Choisit aléatoirement un élément de la liste `L`

Math

`import numpy as np`

`np.atan(x)`	Renvoie $\arctan(x)$
`np.floor(x)`	Renvoie $\lfloor x \rfloor$
`np.factorial(n)`	Renvoie $n!$ si $n \in \mathbb{N}$
`np.sqrt(x)`	Renvoie $\sqrt{x}$ si $x \geq 0$
`np.log(x)`	Renvoie $\ln(x)$ si $x > 0$
`np.exp(x)`	Renvoie e^x

Logique

`a == b`	Teste l'égalité « $a = b$ »
`a != b`	Teste « $a \neq b$ »
`a < b`	Teste « $a < b$ »
`a <= b`	Teste « $a \leq b$ »
`a > b`	Teste « $a > b$ »
`a >= b`	Teste « $a \geq b$ »
`not A`	Renvoie la négation de A
`A and B`	Renvoie « A et B »
`A or B`	Renvoie « A ou B »
`True`	Constante booléenne « Vrai »
`False`	Constante booléenne « Faux »

Matplotlib.pyplot

`import matplotlib.pyplot as plt`

`plt.plot(X,Y,'+-r')`	Génère la courbe des points définis par les listes `X` et `Y` (abscisses et ordonnées) avec les options : • symbole : `'.'` point, `'o'` rond, `'h'` hexagone, `'+'` plus, `'x'` croix, `'*'` étoile, ... • ligne : `'-'` trait plein, `'--'` pointillé, `'-.'` alterné, ... • couleur : `'b'` **bleu** , `'r'` **rouge** , `'g'` **vert** , `'c'` **cyan** , `'m'` **magenta** , `'k'` **noir** , ...
`plt.bar(X,Y)`	Génère l'histogramme des points définis par les listes `X` et `Y` (abscisses et ordonnées)
`plt.axis('equal')`	Rend le repère orthonormé
`plt.xlim(xmin,xmax)`	Fixe les bornes de l'axe des abscisses
`plt.ylim(ymin,ymax)`	Fixe les bornes de l'axe des ordonnées
`plt.show()`	Affiche le graphique

Cette liste est non exhaustive. Les candidats sont libres d'utiliser les commandes de leur choix.

Correction DS 02

Correction Ex.–1

1. Voici la démonstration de l'identité de KOENIG–HUYGHENS :

$$\begin{array}{rcl} \mathbb{V}(X) & = & \mathbb{E}\left((X-\mathbb{E}(X))^2\right) \\ \textbf{[dvpt } (a-b)^2\textbf{]} & = & \mathbb{E}\left(X^2-2X.\mathbb{E}(X)+\mathbb{E}(X)^2\right) \\ \textbf{[linéarité, } \mathbb{E}(\mathbb{1})=1\textbf{]} & = & \mathbb{E}\left(X^2\right)-2\mathbb{E}(X).\mathbb{E}(X)+\mathbb{E}(X)^2 \\ & = & \mathbb{E}\left(X^2\right)-\mathbb{E}(X)^2 \end{array}$$

2. La covariance de deux variables aléatoires X et Y est :

$$\mathbb{C}ov(X,Y)=\mathbb{E}\left((X-\mathbb{E}(X)).(Y-\mathbb{E}(Y))\right)$$

2.a. Encore un calcul basé sur l'identité remarquable $(a+b)^2=a^2+b^2+2ab$ et la linéarité de l'espérance. On a :

$$\begin{array}{rcl} \mathbb{E}((X+Y)^2)=\mathbb{E}(X^2+Y^2+2XY) & = & \mathbb{E}(X^2)+\mathbb{E}(Y^2)+2\mathbb{E}(X.Y) \\ \mathbb{E}((X+Y))^2=(\mathbb{E}(X)+\mathbb{E}(Y))^2 & = & \mathbb{E}(X)^2+\mathbb{E}(Y)^2+2\mathbb{E}(X).\mathbb{E}(Y) \end{array}$$

et finalement, en retranchant, par KOENIG–HUYGHENS pour les variances et définition de la covariance, on reconnait :

$$\mathbb{V}(X+Y)=\mathbb{V}(X)+\mathbb{V}(Y)+2\,\mathbb{C}ov(X,Y).$$

On peut faire le même calcul mais avec trois variables aléatoires en utilisant l'identité remarquable $(a+b+c)^2=a^2+b^2+c^2+2ab+2bc+2ca$. pour obtenir

$$\mathbb{V}(X+Y+Z)=\mathbb{V}(X)+\mathbb{V}(Y)+\mathbb{V}(Z)+2\,\mathbb{C}ov(X,Y)+2\,\mathbb{C}ov(Y,Z)+2\,\mathbb{C}ov(X,Z).$$

2.b. Si X et Y sont indépendantes, $\mathbb{E}(X.Y)=\mathbb{E}(X).\mathbb{E}(Y)$ et donc $\mathbb{C}ov(X,Y)=\mathbb{E}(X.Y)-\mathbb{E}(X).\mathbb{E}(Y)=0$.

2.c. Prenons X à valeurs dans $\{-1,0,+1\}$ uniformément distribuée ($\mathbb{E}(X)=0$) et $Y=X^2$. La v.a. Y est de BERNOULLI vérifiant $\mathbb{P}(Y=1)=\frac{2}{3}$. Calculons $\mathbb{C}ov(X,Y)$:

$$\mathbb{E}(X.Y)=\mathbb{E}(X^3)=(-1)^3.\mathbb{P}(X=-1)+0.\mathbb{P}(X=0)+(+1)^3.\mathbb{P}(X=+1)=0 \text{ et } \mathbb{E}(X).\mathbb{E}(Y)=0.\frac{2}{3}=0$$

Donc $\mathbb{C}ov(X,Y)=0$.

Les v.a. X et Y ne sont pas indépendantes : Comme $\{X=1\}\subset\{Y=1\}$, il vient

$$\mathbb{P}(X=1 \text{ et } Y=1)=\mathbb{P}(X=1)=\frac{1}{3} \text{ et } \mathbb{P}(X=1)\mathbb{P}(Y=1)=\frac{1}{3}.\frac{2}{3}=\frac{2}{9}$$

Ceci montre que $\mathbb{P}(X=1 \text{ et } Y=1)\neq\mathbb{P}(X=1)\mathbb{P}(Y=1)$ et donc que X et Y ne sont pas indépendantes.

3.a. C'est une question de cours (CAUCHY-SCHWARZ) ou presque : Soit la fonction d'une variable réelle f définie par

$$\forall t \in \mathbb{R},\, f(t) = \mathbb{V}(X + t.Y)$$

On a en développant (*cf.* Q.2.a) :

$$\forall t \in \mathbb{R},\, f(t) = \mathbb{V}(X) + 2t\,\mathbb{C}ov(X,Y) + t^2.\mathbb{V}(Y)$$

La fonction f est donc une fonction trinôme, gardant constamment un signe ≥ 0.

— Si $\mathbb{V}(Y) = 0$, f est affine, de coefficient directeur $2\,\mathbb{C}ov(X,Y)$ et si elle ne change pas de signe sur $\mathbb{R}$, c'est que ce coefficient directeur est nul. En ce cas (passablement trivial) :

$$0 = \mathbb{C}ov(X,Y)^2 \leq \mathbb{V}(X).\mathbb{V}(Y) = 0$$

— Si $\mathbb{V}(Y) \neq 0$, f est de degré 2 et, restant de signe constant sur tout $\mathbb{R}$, elle a un discriminant ≤ 0 (si non il y aurait deux racines distinctes et changement de signe au passage de chacune des racines). Ce discriminant vaut $4\,\mathbb{C}ov(X,Y)^2 - 4\mathbb{V}(X).\mathbb{V}(Y)$ et l'inégalité cherchée s'ensuit :

$$\mathbb{C}ov(X,Y)^2 \leq \mathbb{V}(X).\mathbb{V}(Y)$$

3.b. Si $\mathbb{V}(Y) \neq 0$. Le fait qu'il existe $t_0 \in \mathbb{R}$ tel que $f(t_0) = 0$, équivaut au fait que le discriminant précedent est nul, c'est à dire :

$$\mathbb{C}ov(X,Y)^2 = \mathbb{V}(X).\mathbb{V}(Y)$$

Le fait que $\mathbb{V}(X + t_0.Y) = 0$ signifie que $X = -t_0.Y + Cste$ et donc que X est fonction affine de Y.

Réciproquement, s'il existe $\lambda\, mu \in \mathbb{R}$ tel que $X = \lambda.Y + \mu$, on a $f(-\lambda) = 0$

3.c. En assemblant ces deux équivalences (et en discutant des cas $\mathbb{V}(Y) \neq 0$, $\mathbb{V}(X) \neq 0$ et $\mathbb{V}(X) = \mathbb{V}(Y) = 0$), on a l'égalité :

$$\mathbb{C}ov(X,Y)^2 = \mathbb{V}(X).\mathbb{V}(Y)$$

si et seulement si il existe $\lambda, \mu \in \mathbb{R}$ tel que $X = \lambda.Y + \mu$ ou $Y = \lambda.X + \mu$.

4. On définit la fonction f sur $\mathbb{R}^2$ par la formule :

$$\forall (a,b) \in \mathbb{R}^2,\, f(a,b) = \mathbb{E}\left((Z - a.X - b.Y)^2\right)$$

4.a. En développant l'expression définissant $f(a,b)$ comme précédemment (on utilise 2.a, deuxième partie) on a

$$\forall (a,b) \in \mathbb{R},\, f(a,b) = \mathbb{V}(Z) + a^2\mathbb{V}(X) + b^2\mathbb{V}(Y) - 2a.\,\mathbb{C}ov(X,Z) - 2b\,\mathbb{C}ov(Y,Z) + 2ab\,\mathbb{C}ov(X,Y)$$

4.b. La fonction f est une fonction quadratique en le couple de deux variables (a,b) et on a

$$\frac{\partial f}{\partial a}(a,b) = 2a\mathbb{V}(X) - 2\,\mathbb{C}ov(X,Z) + 2b\,\mathbb{C}ov(X,Y) \text{ et } \frac{\partial f}{\partial b}(a,b) = 2b\mathbb{V}(Y) - 2\,\mathbb{C}ov(Y,Z) + 2a\,\mathbb{C}ov(X,Y)$$

Un couple (a,b) est critique pour f si et seulement si $\begin{cases} \frac{\partial f}{\partial a}(a,b) & = & 0 \\ \frac{\partial f}{\partial b}(a,b) & = & 0 \end{cases}$

C'est à dire si et seulement si (a,b) est solution du système linéaire

$$\begin{cases} a\mathbb{V}(X) & + & b\,\mathbb{C}ov(X,Y) & = & \mathbb{C}ov(X,Z) \\ a\,\mathbb{C}ov(X,Y) & + & b\mathbb{V}(Y) & = & \mathbb{C}ov(Y,Z) \end{cases}$$

Le déterminant de ce système est $\Delta = \mathbb{V}(X)\mathbb{V}(Y) - \mathbb{C}ov(X,Y)^2 \neq 0$. Le système est donc de CRAMER, il admet une unique solution non nulle que l'on peut trouver par quelques mouvements de pivot de GAUSS

— $L_2 \leftarrow \mathbb{V}(X)L_2 - \mathbb{C}ov(X,Y)L_1$: $\Delta.b = \mathbb{C}ov(Y,Z)\mathbb{V}(X) - \mathbb{C}ov(X,Y)\,\mathbb{C}ov(X,Z)$ et donc

$$b = \frac{\mathbb{C}ov(Y,Z)\mathbb{V}(X) - \mathbb{C}ov(X,Y)\,\mathbb{C}ov(X,Z)}{\Delta}$$

— $L_1 \leftarrow L_1.V(Y) - \mathbb{C}ov(X,Y)L_2$:

$$a = \frac{\mathbb{C}ov(X,Z)\mathbb{V}(Y) - \mathbb{C}ov(X,Y)\,\mathbb{C}ov(Y,Z)}{\Delta}$$

Correction Ex.–2

1. On $A_1 = 0 + S_1$ et donc

$$\mathbb{P}(A_1 = 1) = \frac{1}{2}, \mathbb{P}(A_1 = 2) = \frac{1}{4} \text{ et } \mathbb{P}(A_1 = 3) = \frac{1}{4}$$

et donc

$$\mathbb{E}(A_1) = \frac{1}{2} + 2.\frac{1}{4} + 3.\frac{1}{4} = \frac{7}{4}$$

et

$$\mathbb{E}(A_1^2) = \frac{1}{2} + 4.\frac{1}{4} + 9.\frac{1}{4} = \frac{15}{4}$$

d'où (par KOENIG–HUYGHENS),

$$\mathbb{V}(A_1) = \frac{15}{4} - \frac{49}{16} = \frac{11}{16}$$

2.a. On a $A_2 = A_1 + S_2$ car l'abscisse à l'issue du 2^e saut est celle à l'issue du premier saut (A_1) à laquelle on a ajouté le nmbre d'unités sautées au deuxième saut (S_2). Comme $A_1 = S_1$, on a bien $A_2 = S_1 + S_2$. Les variables S_1 et S_2 étant supposées indépendantes, on a $\mathbb{P}(S_1 = k, S_2 = \ell) = \mathbb{P}(S_1 = k)\mathbb{P}(S_2 = \ell)$ pour tout couple $(k,\ell) \in \{1,2,3\}^2$. Le tableau suivant donnant la loi de du couple (S_1,S_2) :

$S_1 \backslash S_2$	1	2	3
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
3	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$

La table de la somme $A_2 = S_1 + S_2$ est :

$S_1 \backslash S_2$	1	2	3
1	2	3	4
2	3	4	5
3	4	5	6

et donc A_2 est à valeurs dans $\{2,\dots,6\}$ avec la loi donnée par la table [1] :

k	2	3	4	5	6
$\mathbb{P}(A_2 = k)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

2.b. On a donc

$$\mathbb{E}(A_2) = 2.\frac{1}{4} + 3.\frac{1}{4} + 4.\frac{5}{16} + 5.\frac{1}{8} + 6.\frac{1}{16} = \frac{56}{16} = \frac{7}{2}.$$

3. *cf.* Script 1.

1. Pour chaque k, on repère les cases dans la table de la somme valant ce k et on ajoute les probabilités dans les cases correspondantes de la table de la loi de couple.

4. Pour $n \in \mathbb{N}^*$ fixé, X_n est le nombre de fois où le saut était de une unité dans les sauts numérotés de 1 à n, *i.e.*

$$X_n = \sum_{k=1}^{n} 1\!\!1_{\{S_k=1\}}$$

Comme pour chaque k, $1\!\!1_{\{S_k=1\}}$ est une v.a. de BERNOULLI de paramètre de succès $\frac{1}{2}$, comme les v.a. $1\!\!1_{\{S_k=1\}}, k \in \{1,\dots,n\}$ sont indépendantes (car fonctions de v.a. indépendantes, *cf.* Lemme des coalitions), X_n suit donc une loi binomiale de paramètres n et $\frac{1}{2}$:

$$X_n \sim \mathscr{B}\left(n, \frac{1}{2}\right)$$

Le même raisonnement en partant des identités

$$Y_n = \sum_{k=1}^{n} 1\!\!1_{\{S_k=2\}} \text{ et } Z_n = \sum_{k=1}^{n} 1\!\!1_{\{S_k=3\}}$$

donne

$$Y_n \sim \mathscr{B}\left(n, \frac{1}{4}\right) \text{ et } Z_n \sim \mathscr{B}\left(n, \frac{1}{4}\right)$$

On a

$$X_n + Y_n + Z_n = \sum_{k=1}^{n} 1\!\!1_{\{S_k=1\}} + \sum_{k=1}^{n} 1\!\!1_{\{S_k=2\}} + \sum_{k=1}^{n} 1\!\!1_{\{S_k=3\}} = \sum_{k=1}^{n} (1\!\!1_{\{S_k=1\}} + 1\!\!1_{\{S_k=2\}} + 1\!\!1_{\{S_k=3\}}) = \sum_{k=1}^{n} 1 = n$$

On s'est servi du fait que pour tout k, S_k est à valeurs dans $\{1,2,3\}$ et donc

$$1\!\!1_{\{S_k=1\}} + 1\!\!1_{\{S_k=2\}} + 1\!\!1_{\{S_k=3\}} = 1$$

On a donc

$$X_n + Y_n = n - Z_n$$

et un calcul simple montre que si Z_n suit une loi binomiale de paramètres n et $p = \frac{1}{4}$, $n - Z_n$ suit une loi binomiale de paramètres n et $q = 1 - p = \frac{3}{4}$. En effet, pour $\ell \in \{0,\dots,n\}$,

$$\mathbb{P}(n - Z_n = \ell) = \mathbb{P}(Z_n = n - \ell) = \binom{n}{n-\ell} p^{n-\ell}(1-p)^{n-(n-\ell)} = \binom{n}{\ell} q^{\ell}.(1-q)^{n-\ell}$$

5. Soit $n \in \mathbb{N}^*$. On a [2]

$$\mathbb{V}(X_n) = n.\frac{1}{4},\ \mathbb{V}(Y_n) = n.\frac{1}{4}.\frac{3}{4} = \frac{3n}{16} \text{ et } \mathbb{V}(X_n + Y_n) = n.\frac{3}{4}.\frac{1}{4} = \frac{3n}{16}$$

et en utilisant l'identité

$$\mathbb{V}(X_n + Y_n) = \mathbb{V}(X_n) + \mathbb{V}(Y_n) + 2\,\mathbb{C}ov(X_n, Y_n)$$

il vient

$$\mathbb{C}ov(X_n, Y_n) = \frac{1}{2}.n\left(\frac{3}{16} - \frac{1}{4} - \frac{3}{16}\right) = -\frac{n}{8}$$

Leur covariance étant non nulle, les v.a. X_n et Y_n ne sont pas indépendantes.

2. La variance d'une $\mathscr{B}(n,p)$ est $n.p.(1-p)$.

6.a. On a $A_n = X_n + 2Y_n + 3Z_n$ en comptant simplement le nombre d'unités sautées suivant leur catégorie et donc, par linéarité de l'espérance et valeur de l'espérance d'une v.a. binomiale [3],

$$\mathbb{E}(A_n) = \mathbb{E}(X_n) + 2\mathbb{E}(Y_n) + 3\mathbb{E}(Z_n) = n(\frac{1}{2} + 2\frac{1}{4} + 3\frac{1}{4}) = \frac{7n}{4}.$$

6.b. Sachant que $Z_n = n - X_n - Y_n$, on a donc

$$A_n = 3n - 2X_n - Y_n$$

et donc (règles de calcul sur la variance) :

$$\begin{array}{rcl} \mathbb{V}(A_n) & = & \mathbb{V}(3n - 2X_n - Y_n) \\ & = & \mathbb{V}(-2X_n - Y_n) = \mathbb{V}(+2X_n + Y_n) \\ & = & \mathbb{V}(2X_n) + \mathbb{V}(Y_n) + 2\,\mathbb{C}ov(2X_n, Y_n) = 4\mathbb{V}(X_n) + \mathbb{V}(Y_n) + 4\,\mathbb{C}ov(X_n, Y_n) \\ \textbf{[valeurs connues]} & = & 4.n.\frac{1}{4} + n\frac{3}{16} - 4.n\frac{1}{8} = \frac{11n}{16} \end{array}$$

6.c. De même, par les règles sur la covariance,

$$\begin{array}{rcl} \mathbb{C}ov(A_n, X_n) & = & \mathbb{C}ov(3n - 2X_n - Y_n, X_n) \\ & = & \mathbb{C}ov(3n, X_n) - 2\,\mathbb{C}ov(X_n, X_n) - \mathbb{C}ov(Y_n, X_n) \\ \textbf{[valeurs connues]} & = & 0 - 2n.\frac{1}{4} + n\frac{1}{8} = -\frac{3n}{8} \end{array}$$

7. La sortie graphique obtenue est le graphe de la suite (A_n) pour $n = 0, \ldots, 200$ (complété par des segments intermédiaires).

En supposant que chaque `A[k]` est proche de l'espérance théorique $\mathbb{E}(A_k) = \frac{7}{4}.k$, l'allure de la courbe tracée est donc proche de la droite d'équation $y = \frac{7}{4}.x$. Un graphique explicatif s'obtient en exécutant le script 1.

Listing 1 – python/correction-Q4-rev2023.py

```python
import numpy
import matplotlib.pyplot
#Q.4
def S(N=100):
    s = [0]*(N+1)
    for k in range(N):
        u = numpy.random.rand()
        if u <= 0.5 :
            s[k] = 1
        elif u <= 0.75 :
            s[k] = 2
        else :
            s[k] = 3
    return s
#Q.7
Sauts = S(N=200)
A = [0]*(len(Sauts))
for k in range(1,len(A)):
    A[k] = A[k-1] + Sauts[k]
abscisses = [k for k in range(len(A))]
matplotlib.pyplot.plot(abscisses,A)
```

3. L'espérance d'une v.a. $\mathscr{B}(n,p)$ est $n.p$.

Correction Ex.–3

Partie A

Préliminaires : Fonction logarithme en base 2 et propriété de convexité.

A.1.a. Soit $(x,y) \in (]0,+\infty[)^2$, on a, par la relation fondamentale sur la fonction ln,

$$\log_2(x.y) = \frac{1}{\ln 2}(\ln(x.y)) = \frac{1}{\ln 2}(\ln(x)+\ln(y)) = \log_2(x)+\log_2(y).$$

A.1.b. Soit $\alpha \in \mathbb{R}$, on a, par définition de la fonction puissance,

$$\log_2(2^\alpha) = \frac{\ln(2^\alpha)}{\ln 2} = \frac{\ln(e^{\alpha.\ln 2})}{\ln 2} = \frac{\alpha.\ln 2}{\ln 2} = \alpha.$$

A.1.c. Si $x \in]0,1[$, $\ln x < 0$. Comme $\ln 2 > 0$ (car $2 > 1$, on a donc :

$$\forall x \in]0,1[\,,\, \log_2(x) < 0.$$

A.2. Fonctions convexes sur un intervalle. Soit I un intervalle de $\mathbb{R}$ non vide et non réduit à un point. Une fonction $\phi : I \to \mathbb{R}$, de classe $\mathscr{C}^2$ sur I sera dite *convexe* sur I si

$$\forall x \in I, \phi''(x) \geq 0$$

A.2.a. Soit ϕ définie par $\forall t \in [0,+\infty[\,,\, \phi(t) = 2^t = e^{t.\ln 2}$.

Par composition de l'exponentielle avec une fonction affine, ϕ est $\mathscr{C}^\infty$ (et donc $\mathscr{C}^2$!) sur $I = \mathbb{R}$ et on a

$$\forall t \in \mathbb{R},\, \phi'(t) = \ln 2.e^{t.\ln 2} = \ln 2.2^t$$

et

$$\forall t \in \mathbb{R},\, \phi''(t) = (\ln 2)^2.e^{t.\ln 2} = (\ln 2)^2.2^t > 0$$

La fonction ϕ est donc convexe sur $I = [0,+\infty[$.

A.2.b. Pour $t \in I =]0,+\infty[$, on a $\phi(t) := -\log_2(t) = -\frac{1}{\ln 2}.\ln t$. La fonction ϕ est donc $\mathscr{C}^\infty$ sur I avec

$$\forall t \in]0,+\infty[\,,\, \phi'(t) = -\frac{1}{\ln 2}.\frac{1}{t}$$

et

$$\forall t \in \mathbb{R},\, \phi''(t) = +\frac{1}{\ln 2}.\frac{1}{t^2} > 0$$

La fonction ϕ est donc convexe sur $I =]0,+\infty[$.

A.2.c. On suppose que $\phi : I \to \mathbb{R}$, de classe $\mathscr{C}^2$ sur l'intervalle non trivial I y est convexe.

Soit $x,y \in I$. Posons pour $t \in [0,1]$,

$$g(t) = t.\phi(x) + (1-t).\phi(y) - \phi(t.x + (1-t).y)$$

La fonction $t \in [0,1] \mapsto t.x+(1-t).y$ est affine, à valeurs dans le segment $[x,y] \subset I$ (car I est un intervalle contenant les deux points x et y). La composée

$$t \in [0,1] \mapsto \phi(t.x+(1-t).y)$$

est donc bien définie et de classe $\mathscr{C}^2$ sur $[0,1]$. Par addition avec la fonction affine $t \in [0,1] \mapsto t.\phi(x)+(1-t).\phi(y)$, la fonction g est bien définie, de classe $\mathscr{C}^2$ sur $[0,1]$ avec

$$\forall t \in [0,1]\,,\, g'(t) = \phi(x) - \phi(y) - (x-y).\phi'(t.x+(1-t).y)$$

et

$$\forall t \in [0,1]\,,\, g''(t) = -(x-y)^2.\phi''(t.x+(1-t).y)$$

De ce calcul s'ensuivent les faits suivants :

— la fonction g'' est négative sur $[0,1]$ car $\phi'' \geq 0$;
— la fonction g' est décroissante sur $[0,1]$ avec

$$g'(0) = \phi(x) - \phi(y) - (x-y).\phi'(y) \text{ et } g'(1) = \phi(x) - \phi(y) - (x-y).\phi'(x)$$

— La fonction g vérifie $g(0) = 0$ et $g(1) = 0$.

Donc, par le théorème de ROLLE, il existe $\alpha \in]0,1[$ tel que $g'(\alpha) = 0$. Par décroissance de g',

$$\forall t \in [0,\alpha], g'(t) \geq 0$$

et

$$\forall t \in [\alpha,1], g'(t) \leq 0$$

La fonction g est donc croissante sur $[0,\alpha]$, décroissante sur $[\alpha,1]$, elle est donc partout $\geq 0 = g(0) = g(1)$.

En résumé, on a le tableau de variations suivant

t	0		α		1
$g''(t)$		≤ 0		≤ 0	
$g'(t)$		$\searrow$	0	$\searrow$	
$g'(t)$		≥ 0	0	≤ 0	
$g(t)$	0	$\nearrow$		$\searrow$	0
$g(t)$	0	≥ 0		≥ 0	0

On a donc

$$\forall t \in [0,1], g(t) = t.\phi(x) + (1-t).\phi(y) - \phi(t.x + (1-t).y) \geq 0$$

i.e.

$$\forall t \in [0,1], t.\phi(x) + (1-t).\phi(y) \geq \phi(t.x + (1-t).y)$$

Les nombres x et y étant quelconques dans I, cela montre la proposition demandée. On admet l'énoncé suivant, se déduisant du précédent par une récurrence élaborée :

Si $\phi : I \to \mathbb{R}$, de classe $\mathscr{C}^2$ sur l'intervalle non trivial I y est convexe alors

$$\forall n \in \mathbb{N}^*, \forall (x_1, \ldots, x_n) \in I^n, \forall (t_1, \ldots, t_n) \in ([0,1])^n, \sum_{i=1}^{n} t_i = 1 \Rightarrow \phi\left(\sum_{i=1}^{n} t_i.x_i\right) \leq \sum_{i=1}^{n} t_i.\phi(x_i) \qquad \text{(Cvx)}$$

Partie B

Entropie d'une variable aléatoire réelle discrète.

B.1.a. La variable aléatoire $Y = g(X)$ est donc bien définie, à valeurs dans l'ensemble fini $\{\log_2 \mathbb{P}(X = k), k \in \{0,\ldots,n\}\} \subset]-\infty, 0[$. Par la formule de transfert, comme X est à valeurs dans $\{0,\ldots,n\}$, on a

$$\mathbb{E}(g(X)) = \sum_{k=0}^{n} g(k).\mathbb{P}(X = k) = \sum_{k=0}^{n} \log_2(\mathbb{P}(X = k)).\mathbb{P}(X = k) = -H(X)$$

et donc on a bien $H(X) = -\mathbb{E}(g(X))$.

B.1.b. Comme $g(X)$ est à valeurs négatives, son espérance est négative et $H(X) = -\mathbb{E}(g(X)) \geq 0$.

B.1.c.i. La v.a. X est donc à valeurs dans $\{0,1\}$ avec

$$\mathbb{P}(X = 0) = 1 - p > 0; \mathbb{P}(X = 1) = p > 0$$

La v.a. X est donc à support $\{0,1\}$ et

$H(X) = -\mathbb{P}(X=0).\log_2 \mathbb{P}(X=0) - \mathbb{P}(X=1).\log_2 \mathbb{P}(X=1) = -(1-p)\log_2(1-p) - p.\log_2(p) =: \psi(p)$

B.1.c.ii. On a donc

$$\forall p \in]0,1[,\ -\psi(p) = +(1-p)\log_2(1-p) + p.\log_2(p)$$

La fonction $-\psi$ est, par composition, $\mathscr{C}^2$ sur $I =]0,1[$ et, en dérivant une première fois,

$$\forall p \in]0,1[,\ (-\psi)'(p) = -\log_2(1-p) - \frac{1}{\ln 2} + \log_2(p) + \frac{1}{\ln 2} = -\log_2(1-p) + \log_2(p) = \log_2 \frac{p}{1-p}$$

et donc, en dérivant une deuxième fois,

$$\forall p \in]0,1[,\ (-\psi)''(p) = \frac{1}{\ln 2}.\frac{1}{1-p} + \frac{1}{\ln 2}\frac{1}{p} = \frac{1}{\ln 2}.\frac{1}{p(1-p)}$$

Clairement,

$$\forall p \in]0,1[,\ (-\psi)''(p) \geq 0$$

et donc $-\psi$ est convexe sur $]0,1[$.

B.1.c.iii. On a

$$\forall p \in]0,1[,\ \psi'(p) = -\log_2 \frac{p}{1-p} = \log_2(\frac{1}{p} - 1)$$

Comme pour $p \in]0,1[$,

— $\psi'(p) = 0$ si et seulement si $\frac{1}{p} - 1 = 1$ *i.e.* $p = \frac{1}{2}$;

— $\psi'(p) < 0$ si et seulement si $\frac{1}{p} - 1 < 1$ *i.e.* $p > \frac{1}{2}$;

— $\psi'(p) > 0$ si et seulement si $\frac{1}{p} - 1 > 1$ *i.e.* $p < \frac{1}{2}$;

On voit que la fonction ψ est strictement décroissante sur $\left]0, \frac{1}{2}\right]$, strictement croissante sur $\left[\frac{1}{2}, 1\right[$ et donc atteint son maximum sur $]0,1[$ uniquement au point $p_0 = \frac{1}{2}$. Le maximum vaut $-\log_2 \frac{1}{2} = +1$.

B.1.d. On a les probabilités

$$\mathbb{P}(X=0) = \frac{1}{2}; \mathbb{P}(X=1) = \frac{1}{4}; \mathbb{P}(X=2) = \mathbb{P}(X=3) = \frac{1}{8}.$$

On a donc (on utilise $\log_2 2^{-k} = -k$)

$$\begin{aligned} H(X) &= -\frac{1}{2}.\log_2 \frac{1}{2} - \frac{1}{4}.\log_2 \frac{1}{4} - \frac{1}{8}.\log_2 \frac{1}{8} - \frac{1}{8}.\log_2 \frac{1}{8} \\ &= +\frac{1}{2} + 2.\frac{1}{4} + 3.\frac{1}{8} + 3.\frac{1}{8} = \frac{14}{8} = \frac{7}{4} \end{aligned}$$

B.2.

Listing 2 – python/entropie.py

```python
import numpy as np

def Entropie(P):
    """
    Retourne l'entropie de X v.a. à support {0,..,n} dont la loi est décrite
    par le vecteur P de longueur n+1
    """
    h=0
    for k in range(len(P)):
        h += P[k]*np.log(P[k])
    return -h/np.log(2)
```

B.3. On a, du fait que X est de loi à support $\{x_0, x_1, \ldots, x_n\}$, que

$$\forall i \in \{0, \ldots, n\}, \, p_i = \mathbb{P}(X = x_i) > 0.$$

Par ailleurs, $\sum_{i=0}^{n} p_i = 1$. Si l'un des p_i valait 1, tous les autres (il y en a au moins car $n+1 \geq 2$) seraient donc nuls, ce qui est exclus par la définition de support. Donc

$$\forall i, 0 \leq i \leq n, \, p_i < 1.$$

Si I est un intervalle de $\mathbb{R}$ contenant les nombres x_i et $\phi : I \to \mathbb{R}$ est une fonction convexe sur I.

On a $\mathbb{E}(X) = \sum_{i=0}^{n} x_i . p_i$. $\mathbb{E}(X)$ est donc une moyenne pondérée par des coefficients positifs des nombres x_i. En tant que telle $\min_i(x_i) \leq \mathbb{E}(X) \leq \max_i(x_i)$ et, I étant un intervalle contenant $\min_i(x_i)$ et $\max_i(x_i)$, il contient $\mathbb{E}(X)$.

On a donc

$$\mathbb{E}(X) = \sum_{i=0}^{n} x_i . p_i$$

et, par la formule de transfert,

$$\mathbb{E}(\phi(X)) = \sum_{i=0}^{n} \phi(x_i) . p_i$$

Comme les p_i sont positifs et que $\sum_i p_i = 1$, et ϕ est convexe sur I, l'inégalité (Cvx) se réécrit (en prenant $\forall i, t_i = p_i$) :

$$\phi\left(\mathbb{E}(X)\right) \leq \mathbb{E}(\phi(X))$$

B.4. Soit X une variable aléatoire de loi à support $\{0, \ldots, n\}$. On pose, $p_k = \mathbb{P}(X = k)$.

B.4.a. Considérons les nombres dans $I =]0, +\infty[$, $y_0 = \frac{1}{(n+1)p_0}$, ..., $y_n = \frac{1}{(n+1)p_n}$ et $t_0 = p_0$,... $t_n = p_n$ et la fonction $\phi : t \in I \to \mathbb{R}$ définie par $\phi(t) = -\log_2 t$.

On a vu en A.2.b que la fonction ϕ est convexe sur I et on a donc, par l'inégalité (Cvx) appliquée aux y_i, t_i que

$$\phi\left(\sum_i t_i . y_i\right) \leq \sum_i t_i . \phi(y_i)$$

ce qui s'écrit

$$-\log_2\left(\sum_i t_i . y_i\right) \leq \sum_i t_i . (-\log_2(y_i))$$

i.e. (en passant à l'opposé)

$$\log_2\left(\sum_i p_i . \frac{1}{(n+1)p_i}\right) \geq \sum_i p_i . \log_2 \frac{1}{(n+1)p_i}$$

On a $\sum_i p_i . \frac{1}{(n+1)p_i} = \sum_i \frac{1}{(n+1)} = 1$ et donc

$$\log_2\left(\sum_i p_i . \frac{1}{(n+1)p_i}\right) = 0$$

et finalement,

$$\sum_i p_i . \log_2 \frac{1}{(n+1)p_i} \leq 0$$

B.4.b. On a, en utilisant $\log_2((n+1)p_k) = \log_2(n+1) + \log_2 p_k$, que

$$\sum_{k=0}^{n} p_k \log_2[(n+1)p_k] = \sum_{k=0}^{n} p_k \log_2(n+1) + \sum_{k=0}^{n} p_k \log_2 p_k = \log_2(n+1) - H(X)$$

B.4.c. Comme
$$\sum_k p_k.\log_2[(n+1)p_i] = -\sum_k p_k.\log_2 \frac{1}{(n+1)p_k} \geq 0$$
On a donc
$$\log_2(n+1) - H(X) \geq 0$$
et donc $H(X) \leq \log_2(n+1)$.

B.4.d. On suppose que X suit la loi uniforme sur $\{0,\ldots,n\}$, on a $p_k = \frac{1}{n+1}$ et donc
$$H(X) = -\sum_{k=0}^{n} \frac{1}{n+1} \log_2 \frac{1}{n+1} = \log_2(n+1)$$
On en déduit que parmi les v.a. dont la loi a support $\{0,\ldots,n\}$, les v.a. uniformes ont l'entropie maximale.

B.5.a. On décompose l'évenement $\{X = Y\}$ suivant les valeurs de X : On a donc
$$\{X = Y\} = \cup_{k=0}^{n} \{X = Y \text{ et } X = k\}$$
et donc, par réécriture logique,
$$\{X = Y\} = \cup_{k=0}^{n} \{X = k \text{ et } Y = k\}$$
Il s'agit d'une union d'événements deux à deux incompatibles et donc
$$\mathbb{P}(X = Y) = \sum_{k=0}^{n} \mathbb{P}(X = k \text{ et } Y = k)$$
Comme X et Y sont indépendantes, $\forall k$, $\mathbb{P}(X = k \text{ et } Y = k) = \mathbb{P}(X = k).\mathbb{P}(Y = k)$ et comme X et Y ont même loi, $\forall k$, $\mathbb{P}(X = k \text{ et } Y = k) = \mathbb{P}(X = k)^2$ et finalement
$$\mathbb{P}(X = Y) = \sum_{k=0}^{n} \mathbb{P}(X = k)^2.$$
B.5.b. On pose $v(k) = \mathbb{P}(X = k)$ pour tout k élément de $\{0,\ldots,n\}$.

Par A.2.a, la fonction $\phi = t \mapsto 2^t$ est convexe sur $\mathbb{R}$ et donc d'après l'inégalité de JENSEN, question B.3, appliquée à la v.a. $Y = \log_2 v(X)$ (on remarque aussi que $\phi(\log_2 v(X)) = v(X)$ par A.1.b), on a
$$\phi(\mathbb{E}(\log_2 v(X)) \leq \mathbb{E}(\phi(\log_2 v(X))$$
et ceci se réécrit
$$2^{\mathbb{E}(\log_2 v(X))} \leq \mathbb{E}(2^{\log_2 v(X)}) = \mathbb{E}(v(X)).$$
B.5.c. On a d'une part
$$\mathbb{E}(\log_2 v(X)) = -H(X)$$
et d'autre part, par la formule de transfert,
$$\mathbb{E}(v(X)) = \sum_{k=0}^{n} v(k).\mathbb{P}(X = k) = \sum_{k=0}^{n} \mathbb{P}(X = k).\mathbb{P}(X = k) = \sum_{k=0}^{n} \mathbb{P}(X = k)^2$$
et donc la formule précédente se réécrit
$$2^{-H(X)} \leq \sum_{k=0}^{n} \mathbb{P}(X = k)^2 = \mathbb{P}(X = Y).$$
B.5.d. Si X est uniforme sur $\{0,\ldots,n\}$, d'une part $H(X) = \log_2(n+1)$ et donc $2^{-H(X)} = \frac{1}{n+1}$ et d'autre part
$$\sum_{k=0}^{n} \mathbb{P}(X = k)^2 = \sum_{k=0}^{n} \frac{1}{(n+1)^2} = \frac{1}{n+1}$$
Il s'agit donc d'un exemple de loi où l'inégalité précédente est une égalité.