Université Paris 8 Année 2020–2021

Master Mathématiques et applications

Théorie de l'information – Partiel – Solutions

sujet du 08 janvier 2021
corrigé le 5 février 2021

Exercice 1. Entropie.

Soient X et Y deux variables aléatoires discrètes à valeurs dans $\mathbb{R}$, définies sur un même espace probabilisé. On définit ensuite $Z = X + Y$.

Question 1.– Démontrer que $H(Z) \leq H(X) + H(Y)$.

Question 2.– Démontrer que $H(Z \mid X) = H(Y \mid X)$.

Question 3.– Démontrer que si X et Y sont indépendantes, alors on a

$$H(X) \leq H(Z) \quad \text{et} \quad H(Y) \leq H(Z)\,.$$

Question 4.– Trouver un exemple de variables X et Y pour lesquelles on a simultanément

$$H(Z) < H(X) \quad \text{et} \quad H(Z) < H(Y)\,.$$

Solutions de l'Exercice 1.

Solution Q1. On a

$$H(Z) = H(X + Y) \leq H(X, Y) \leq H(X) + H(Y)$$

où la première inégalité est dûe au principe de non-création d'information.

Solution Q2. D'une part, on a

$$H(Z \mid X) \leq H(X, Y \mid X) \leq H(Y \mid X) + H(X \mid X) = H(Y \mid X).$$

D'autre part (le raisonnement est similaire),

$$H(Y \mid X) \leq H(Z - X \mid X) \leq H(Z \mid X) + H(X \mid X) = H(Z \mid X).$$

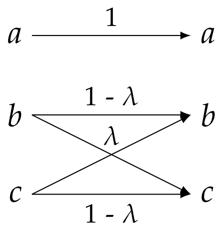
Solution Q3. Si X et Y dont indépendantes, alors $H(X \mid Y) = H(X)$ et $H(Y \mid X) = H(Y)$. D'après la question 2 (où X et Y peuvent jouer des rôles symétriques), on obtient donc

$$H(Z) \geq H(Z \mid X) = H(Y \mid X) = H(Y) \quad \text{et} \quad H(Z) \geq H(Z \mid Y) = H(X \mid Y) = H(X)\,.$$

Solution Q4. On peut prendre par exemple $X = -Y = c$ où c est un constante, de sorte que $Z = c$ et $H(Z) = 0$. Si $H(X) > 0$, on a alors également $H(Y) = H(X) > 0$.

Exercice 2. Capacité d'un canal.

On considère le canal représenté par le diagramme suivant. L'entrée (représentée par la variable X) et la sortie (représentée par la variable Y) sont définies sur l'alphabet $\{a, b, c\}$.



Question 1.– Donner la matrice de transition du canal.

Question 2.– On note $x := \mathbb{P}(X = a)$. Démontrer que

$$H(Y \mid X) = (1 - x)\, h(\lambda)$$

où $h(\lambda) = \lambda \log_2 \frac{1}{\lambda} + (1 - \lambda) \log_2 \frac{1}{1-\lambda}$ est l'entropie binaire.

Question 3.– On note $y := \frac{\mathbb{P}(X=b)}{\mathbb{P}(X=b)+\mathbb{P}(X=c)}$. Démontrer que

$$H(Y) = h(x) + (1 - x)\, h((1 - \lambda)y + \lambda(1 - y))\,.$$

Question 4.– En déduire que l'information mutuelle entre X et Y s'exprime comme

$$I(X\,; Y) = h(x) + (1 - x) f(y)\,.$$

où $f(y) := h((1 - \lambda)y + \lambda(1 - y)) - h(\lambda)$.

Question 5.– Dans cette question uniquement, on suppose que $\lambda = \frac{1}{2}$. Que vaut la capacité du canal dans ce cas ? Interpréter.

Question 6.– Dans cette question uniquement, on suppose que $\lambda = 0$. Que vaut la capacité du canal dans ce cas ? Interpréter.

Question 7.– **[plus difficile]** Dans cette question, le paramètre λ est de nouveau un réel quelconque dans $[0, 1]$. Pour quelle valeur de y l'information mutuelle est-elle maximale ? En déduire la capacité du canal.

Solutions de l'Exercice 2.

Solution Q1. La matrice de transition est

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \lambda & \lambda \\ 0 & \lambda & 1 - \lambda \end{pmatrix}$$

Solution Q2. On a

$$\begin{aligned} H(Y \mid X) &= xH(Y \mid X = a) + \mathbb{P}(X = b)H(Y \mid X = b) + \mathbb{P}(X = c)H(Y \mid X = c) \\ &= x \times 0 + \mathbb{P}(X = b)h(\lambda) + \mathbb{P}(X = c)h(\lambda) \\ &= (1 - x)h(\lambda)\,. \end{aligned}$$

Solution Q3. On note $t = (1 - \lambda)y + \lambda(1 - y)$. La distribution de Y est donc

$$(x, (1 - x)t, (1 - x)(1 - t))\,.$$

Les calculs donnent ensuite (le log est en base 2)

$$\begin{aligned}
H(Y) &= -x\log x - (1-x)t\log((1-x)t) - (1-x)(1-t)\log((1-x)(1-t)) \\
&= -x\log x - (1-x)t\log(1-x) - (1-x)t\log(t) - (1-x)(1-t)\log(1-x) - (1-x)(1-t)\log(1-t) \\
&= -x\log x - (1-x)\log(1-x) - (1-x)(t\log(t) + (1-t)\log(1-t)) \\
&= h(x) + (1-x)h(t)\,.
\end{aligned}$$

Solution Q4. Il résulte des questions précedentes que

$$I(X;Y) = H(Y) - H(Y \mid X) = h(x) + (1-x)h((1-\lambda)y + \lambda(1-y)) - (1-x)h(\lambda) = h(x) + (1-x)f(y)\,.$$

Solution Q5. Si $\lambda = 1/2$, alors on observe que $f(y) = h(1/2) - h(1/2) = 0$ pour tout y. Donc

$$I(X;Y) = h(x)$$

est maximale lorsque $x = 1/2$. La capacité du canal vaut donc 1 dans ce cas. En effet, le canal transmet de manière sûre un bit d'information : le fait que $X = a$ ou non.

Solution Q6. Si $\lambda = 0$, alors $H(Y \mid X) = 0$ et il suffit maintenant de maximiser $H(Y)$. Lorsque X est uniforme, Y l'est aussi et on obtient la valeur maximale $H(Y) = \log_2 3$, qui est donc également la capacité du canal. L'interprétation est claire : pour $\lambda = 0$, le canal transmet parfaitement l'information de la source donc sa capacité est maximale.

Solution Q7. Les grandeurs x et y peuvent varier indépendamment, et $I(X;Y) = h(x) + (1-x)f(y)$. Si $1-x > 0$, la valeur de I est maximale lorsque f est maximale. Au vu de la forme de f, on doit maximiser une entropie binaire, ce qui est le cas lorsque

$$(1-\lambda)y + \lambda(1-y) = \frac{1}{2} \iff y = \frac{1}{2}\,.$$

Notons qu'on a alors $f(y) = 1 - h(\lambda)$. Pour maximiser $I(X;Y)$, il suffit donc ensuite de maximiser

$$\phi : x \mapsto h(x) + (1-x)(1-h(\lambda))\,.$$

On a

$$\phi'(x) = h'(x) - (1-h(\lambda)) = \log(\frac{1-x}{x}) - (1-h(\lambda))$$

qui s'annule donc en $x_\lambda = \frac{1}{2^{1-h(\lambda)}+1}$. Il reste donc à calculer $h(x_\lambda) + (1-x_\lambda)(1-h(\lambda))$. On note d'abord que

$$h(x_\lambda) = x_\lambda \log_2(2^{1-h(\lambda)}+1) + (1-x_\lambda)\log_2(\frac{2^{1-h(\lambda)}+1}{2^{1-h(\lambda)}}) = \log_2(2^{1-h(\lambda)}+1) - (1-x_\lambda)\log_2(2^{1-h(\lambda)})\,.$$

Ainsi, la capacité vaut

$$\log_2(2^{1-h(\lambda)}+1) - (1-x_\lambda)\log_2(2^{1-h(\lambda)}) + (1-x_\lambda)(1-h(\lambda)) = \log_2(2^{1-h(\lambda)}+1)\,.$$

Exercice 3. Encodage de Liv-Zempel.

On considère le code de Liv-Zempel, comme vu en cours, défini sur l'alphabet binaire $\{a, b\}$.

Question 1.– Encoder le mot $\boldsymbol{m} = aabbbabaabaabbbaab$ avec le code de Liv-Zempel.

Question 2.– Le mot suivant est-il un encodage de Liv-Zempel valide (justifier la réponse) ?

$$(0,a), (0,b), (1,a), (3,b), (3,a), (2,b), (3,a), (5,a)$$

Solutions de l'Exercice 3.

Solution Q1. On obtient $[(0,a), (1,b), (0,b), (3,a), (4,a), (5,b), (3,b), (1,a)]$. Puis il reste la lettre b en fin de mot, que l'on peut coder de différentes manières, la plus usuelle étant $(3, \varnothing)$.

Solution Q2. Non, car il y a 2 fois le même terme $(3, a)$, ce qui contredit l'unicité dans le dictionnaire.

Exercice 4. Code de Huffman.

La suite de Fibonacci $(F_n)_{n\geq 1}$ est donnée par :

$$F_1 = 1, \quad F_2 = 1, \quad F_n = F_{n-1} + F_{n-2} \text{ pour } n \geq 3\,.$$

Par exemple, on a $F_3 = 2$, $F_4 = 3$ et $F_9 = 34$. On note $S_n = \sum_{i=1}^n F_i$.

Pour $n \geq 2$ fixé, on considère une source $X^{(n)}$ définie sur l'alphabet $\{x_1, \ldots, x_n\}$, et telle que

$$\mathbb{P}(X = x_i) = \frac{F_i}{S_n} \text{ pour tout } i \in \{1, \ldots, n\}$$

Question 1.– Donner la distribution de probabilité de la source pour $n = 6$.

Question 2.– Construire l'arbre binaire du code de Huffman associé à la source $X^{(6)}$.

On souhaite maintenant étudier la forme de l'arbre de Huffman pour $n \geq 2$ quelconque.

Question 3.– Démontrer que $S_n = F_{n+2} - 1$ pour tout $n \geq 1$.

Question 4.– En déduire la forme de l'arbre de Huffman associé. Le résultat devra être démontré formellement.

Question 5.– **[Plus difficile]** Démontrer qu'asymptotiquement, la longueur moyenne du code de Huffman de source $X^{(n)}$ est finie. On pourra s'aider (sans les démontrer) des formules ci-dessous, valables pour tout $n \geq 1$:

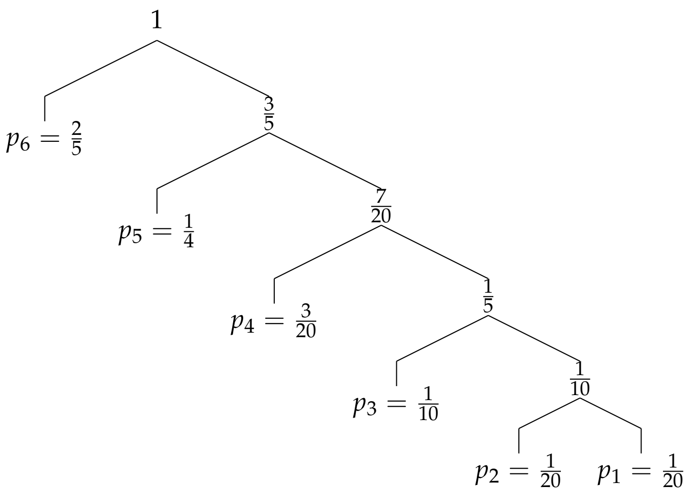
$$\phi^{n-2} \leq F_n \leq \phi^{n-1} \text{ où } \phi := \frac{1+\sqrt{5}}{2} > 1 \qquad \text{et} \qquad \sum_{i=1}^n i x^i = \frac{x}{(x-1)^2}\Big((n-1)x^n - n x^{n-1} + 1\Big)\,.$$

Solutions de l'Exercice 4.

Solution Q1. On obtient

$$(p_1, p_2, p_3, p_4, p_5, p_6) = \left(\tfrac{1}{20}, \tfrac{1}{20}, \tfrac{1}{10}, \tfrac{3}{20}, \tfrac{1}{4}, \tfrac{2}{5}\right)$$

Solution Q2.



Solution Q3. Par récurrence. C'est vrai pour $n = 1$, car $S_1 = 1 - F_3 - 1$. Par ailleurs, $S_{n+1} = S_n + F_{n+1} = F_{n+2} + F_{n+1} - 1 = F_{n+3} - 1$.

Solution Q4. On montre que l'arbre est de hauteur $n-1$. Pour cela, on montre qu'à l'étape i, la probabilité p_{i+1} est choisie avec la probabilité correspondant au sous-arbre déjà construit (ce sous-arbre étant p_1, par convention, à la première étape). D'abord, p_1 et p_2 sont choisies à la première étape, et remplacés par $p'_2 = \frac{S_2}{S_n} = \frac{F_3-1}{S_n} < p_3$. Ainsi, à l'étape 2 sont choisies p_3 et p'_2, remplacées par $p'_3 = \frac{S_3}{S_n}$. Par induction, à l'étape i, p_{i+1} et p'_i sont choisies et remplacées par $p'_{i+1} = \frac{S_{i+1}}{S_n} = \frac{F_{i+3}-1}{S_n} < p_{i+3}$.

Solution Q5. Plusieurs méthodes. On peut par exemple démontrer que l'entropie de $X^{(n)}$ est finie, et utiliser le théorème du cours affirmant que le code de Huffman est de longueur moyenne strictement inférieure à $H(X^{(n)})+1$. Ici, on a

$$H(X^{(n)}) = \sum_{i=1}^{n} \frac{F_i}{S_n} \log_2 \frac{F_i}{S_n} = \frac{1}{S_n}\left(\sum_{i=1}^{n} F_i \log_2 F_i\right) - \log_2(S_n)$$

Comme $F_n \leq \phi^{n-1}$ et $S_n = F_{n+2} - 1 \geq \phi^n - 1$, on obtient :

$$H(X^{(n)}) \leq \left(\frac{1}{\phi^n - 1}\sum_{i=1}^{n}(n-1)\phi^{n-1}\log_2\phi\right) - \log_2(\phi^n - 1)$$

Puis,

$$H(X^{(n)}) \leq \left(\frac{\log_2\phi}{\phi^n-1} \cdot \frac{\phi}{(\phi-1)^2} \cdot ((n-2)\phi^{n-1} - (n-1)\phi^{n-2} + 1)\right) - \log_2(\phi^n - 1)$$

Comme $\phi^2 = \phi + 1$, on a alors

$$H(X^{(n)}) \leq \left(\frac{\log_2\phi}{\phi^n-1} \cdot \frac{\phi}{(\phi-1)^2} \cdot (\phi^{n-3}(n-2-\phi)+1)\right) - n\log_2\phi + \log_2\frac{1}{1-\phi^{-n}}\,.$$

Cette dernière quantité a pour développement asymptotique

$$\frac{\phi\log_2\phi}{(\phi-1)^2}\frac{\phi^{n-3}}{\phi^n-1}(n-2-\phi+o(1)) - n\log_2\phi + o(1) = n\log_2\phi\cdot\left(\frac{\phi}{(\phi-1)^2}\cdot\frac{1}{\phi^3} - 1\right) + O(1)$$

et il reste à noter que, comme $\phi^2 = \phi + 1$, on a

$$\frac{\phi}{(\phi-1)^2}\cdot\frac{1}{\phi^3} = \frac{1}{(\phi-1)^2(\phi+1)} = \frac{1}{(\phi-1)\phi} = 1\,.$$
