Université Paris 8 Année 2020–2021

Master Mathématiques et applications

Théorie de l'information – Partiel (2h)

08 janvier 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

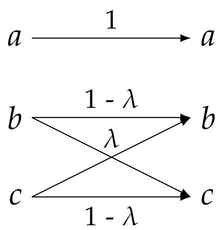
$$H(X) \leq H(Z) \quad \text{et} \quad H(Y) \leq H(Z)\,.$$

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

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

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.

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) ?

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

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 $\{x_1,\dots,x_n\}$, et telle que

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

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 êtré 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 $n \geq 1$:

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