

Théorie de l'information : DS du 26 octobre 2021

*Master Sciences et Technologies, mention Mathématiques ou Informatique,
parcours Cryptologie et Sécurité informatique*

Responsable : Gilles Zémor

Durée : 1h30. Sans document. Les exercices sont indépendants.

– EXERCICE 1.

- a) Soit $X = (X_1, X_2, X_3, X_4, X_5)$ la variable aléatoire de loi uniforme et prenant ses valeurs parmi les cinq décalés circulaires du quintuplet binaire (11100). Calculer $H(X_1)$, $H(X_1|X_2)$ et $I(X_1, X_2)$.
- b) Soit $\mathcal{X} \subset \{0, 1\}^7$ l'ensemble à 8 éléments constitué du septuplet nul et des 7 décalés circulaires du septuplet (1011100). On considère

$$X = (X_1, X_2, \dots, X_7)$$

de loi uniforme dans \mathcal{X} . Calculer $H(X_1)$, $H(X_1, X_2)$, $H(X_1, X_2, X_3)$ ainsi que $H(X_1|X_2)$ et $H(X_1|X_2, X_3)$. Que vaut $H(X_1|X_2, X_3, X_4)$?

– EXERCICE 2. Soit X une variable prenant m valeurs et de loi $p = (p_1, \dots, p_m)$. Soit ℓ_1, \dots, ℓ_m la distribution des longueurs d'un code préfixe tel que

$$\sum_{i=1}^m \frac{1}{2^{\ell_i}} = 1.$$

Soit $\bar{\ell}$ la longueur moyenne de l'encodage de X par ce code. Exprimer $\bar{\ell} - H(X)$ comme une divergence de Kullback.

– EXERCICE 3. On dispose d'un lot de 9 pièces d'apparences identiques, numérotées de 1 à 9. L'une pèse 3 grammes et les 8 autres 5 grammes. On dispose d'une balance à plateau unique qui donne le poids de ce qu'on pose sur le plateau. On souhaite découvrir le numéro de la pièce de 3 grammes. Calculer à l'aide d'un arbre de Huffman le nombre moyen de pesées nécessaire pour découvrir la pièce.

– EXERCICE 4. Soit (p_1, p_2, p_3, p_4) une loi de probabilité. Quitte à permuter les indices on adoptera toujours la convention $p_1 \geq p_2 \geq p_3 \geq p_4$.

- a) Trouver toutes les valeurs possibles de (p_3, p_4) pour lesquelles le code $\{00, 01, 10, 11\}$ est un code de Huffman pour une certaine loi (p_1, p_2, p_3, p_4) .
- b) Trouver toutes les valeurs possibles de (p_1, p_2) pour lesquelles le code $\{00, 01, 10, 11\}$ est un code de Huffman pour une certaine loi (p_1, p_2, p_3, p_4) .

– EXERCICE 5. Soit X_1 une variable de Bernoulli de loi uniforme $P(X_1 = 0) = P(X_1 = 1) = 1/2$. On crée ensuite une variable de Bernoulli X_2 de la manière suivante. On tire à pile ou face : si c'est pile on prend $X_2 = X_1$, et si c'est face on prend X_2 uniforme indépendante de X_1 . Puis on continue ainsi, c'est-à-dire qu'on a deux suites de variables de Bernoulli, X_1, \dots, X_n, \dots et Z_1, \dots, Z_n, \dots , où (Z_1, \dots, Z_n) est uniforme dans $\{0, 1\}^n$, et Z_n indépendante de (X_1, \dots, X_n) , et où

- Si $Z_n = 0$, alors $X_{n+1} = X_n$,
- Si $Z_n = 1$, alors X_{n+1} est choisi uniforme et indépendante de (X_1, \dots, X_n) , c'est à dire que pour tout $(x_1, \dots, x_n) \in \{0, 1\}^n$,

$$P(X_{n+1} = 0 | Z_n = 1, X_1 = x_1, \dots, X_n = x_n) = \frac{1}{2}.$$

- a) Calculer la loi de (X_1, X_2) et en déduire $H(X_1, X_2)$. Faire de même avec (X_1, X_2, X_3)
- b) Montrer que $H(Z_n | X_1, \dots, X_{n+1}) = \frac{3}{4}h(1/3)$.
- c) Calculer $H(X_{n+1} | X_1, \dots, X_n, Z_n)$, puis $H(X_1, \dots, X_n, X_{n+1}, Z_n)$ en fonction de $H(X_1, \dots, X_n)$.
- d) En déduire $H(X_1, \dots, X_n, X_{n+1})$ en fonction de $H(X_1, \dots, X_n)$. En déduire par récurrence la valeur de $H(X_1, \dots, X_n)$.
- e) Que vaut la limite de $\frac{1}{n}H(X_1, \dots, X_n)$? Comment peut-on utiliser l'algorithme de Huffman pour transformer X_1, \dots, X_n en une suite s_n de symboles binaires dont la longueur ℓ_n est telle que

$$\frac{\ell_n}{n} - \frac{1}{n}H(X_1, \dots, X_n) \rightarrow 0$$

quand n tend vers l'infini ?