

## Théorie de l'information : DS du 20 octobre 2014

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

*Responsable* : Gilles Zémor

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

– EXERCICE 1. Soit  $p$  une loi  $p = (p_1, p_2, \dots, p_m)$ . Soit  $u$  la loi uniforme sur  $m$  objets. Montrer que  $H(p) = \log_2 m - D(p||u)$ .

– EXERCICE 2. On considère une variable aléatoire  $X$  de loi uniforme dans l'ensemble  $\{1, 2, 3, 4\}$ . Puis, après avoir observé la valeur de  $X$ , on crée une variable  $Y$ , de loi uniforme dans l'ensemble des entiers  $i$  avec  $X \leq i \leq 4$ . Calculer  $H(Y|X)$ ,  $H(X|Y)$ ,  $H(X, Y)$ ,  $H(X)$  et  $H(Y)$ .

– EXERCICE 3.

a) Construire un code préfixe avec comme distribution des longueurs

1, 3, 3, 3, 5, 5.

b) Quelles sont toutes les distributions des longueurs possibles d'un code préfixe associé à une loi sur 5 symboles ?

– EXERCICE 4. Soit  $p = (p_1, \dots, p_m)$  une loi où on a supposé les  $p_i$  ordonnés  $p_1 \geq p_2 \geq \dots \geq p_m$ .

a) Montrer que si  $p_1 < 1/3$ , alors tous les mots d'un code de Huffman associé à  $p$  sont de longueur au moins 2.

b) Donner un exemple de loi  $p_1 \geq \dots \geq p_m$  telle que  $p_1 > 1/3$  et le code de Huffman associé à  $p$  a tous ses mots de longueur au moins 2.

– EXERCICE 5. Soit  $X$  une variable de loi uniforme prenant ses valeurs dans l'ensemble  $\mathcal{X} = \{000, 001, 010, 100\}$ . Une variable  $Y$  est créée à partir de  $X = X_1X_2X_3$  on supprimant aléatoirement un symbole de  $X$ , pour donner  $X_2X_3$  avec probabilité  $1/3$ ,  $X_1X_3$  avec probabilité  $1/3$ , ou  $X_1X_2$  avec probabilité  $1/3$ . La variable  $Y$  prend donc ses valeurs dans l'ensemble  $\mathcal{Y} = \{00, 01, 10\}$ .

Calculer  $H(Y)$  et l'information mutuelle  $I(X, Y)$ . Que vaut  $H(X|Y)$  ?

– EXERCICE 6. On sait qu'une rare bactérie nocive se trouve dans le vin provenant d'un certain vignoble qu'on cherche à localiser. Des études préliminaires restreignent l'ensemble des vignobles possibles à un ensemble de six vignobles. Pour

ces six vignobles, numérotés de 1 à 6, la probabilité d'être le porteur de la bactérie est donnée par le 6-uple  $(p_1, p_2, \dots, p_6) = (8/23, 6/23, 4/23, 2/23, 2/23, 1/23)$ . On applique maintenant des tests à des échantillons consistant en un mélange de vins de différents vignobles.

- a) Donner une borne inférieure sur l'espérance du nombre de tests nécessaire pour déterminer le vignoble infecté.
- b) Par quel mélange de vins faut-il commencer le premier test afin de minimiser le nombre de tests ?