université
PARIS DIDEROT
PARIS 7

M1 - THÉORIE DE L'INFORMATION - 2019-2020

CC DU VENDREDI 09 NOVEMBRE (1 HEURE)

Exercice 1.

Soient X et Y des variables aléatoires à valeurs dans l'ensemble $\{1,2,3,4,5\}$ de lois respectives p et q, qui sont données dans le tableau suivant

x	$p(x)$	$q(y)$	$C(x)$	$D(x)$
1	1/2	1/2	0	0
2	1/4	1/8	10	100
3	1/8	1/8	110	101
4	1/16	1/8	1110	110
5	1/16	1/8	1111	111

1) Calculer les entropies $H(X), H(Y)$, ainsi que les divergences $D(p||q), D(q||p)$. On utilisera des logarithmes en base 2.
2) Les deux dernières colonnes du tableau décrivent respectivement les codes binaires C et D. Démontrer que la longueur moyenne de $C(X)$ est égale à $H(X)$.
3) En déduire que le code de C est optimal relativement à la loi p. Vérifier aussi que le code D est optimal relativement à la loi q.
4) Calculer la longueur moyenne de $D(X)$. Que pouvez-vous dire de la différence $\mathbb{E}(\ell(D(X))) - H(X)$?

Correction.
1) $H(X) = \log(2)/2 + \log(4)/4 + \log(8)/8 + \log(16)/16 + \log(16)/16 = \frac{15}{8}$. De même, $H(Y) = 2$, et $D(p||q) = D(q||p) = \frac{1}{8}$.
2) Ici aussi c'est un simple calcul: $\mathbb{E}(\ell(C(X))) = \sum_x p(x)\ell(C(x)) = 1/2 + 2/4 + 3/8 + 4/16 + 4/16 = 15/8$.
3) Le code C est préfixe donc uniquement décodable. Tout code uniquement décodable à une longueur moyenne bornée inférieurement par l'entropie (Shannon). Donc C est optimal pour p. De même, on peut calculer que $\mathbb{E}(\ell(D(Y))) = \sum_y q(y)\ell(D(y)) = 2 = H(Y)$, et par les mêmes arguments déduire que D est optimal pour q.
4) $\mathbb{E}(\ell(D(X))) = \sum_x p(x)\ell(D(x)) = 2 > H(X)$. Comme la différence est non nulle alors que le code C atteint exactement l'entropie, on en déduit que D n'est pas optimal pour p.

Exercice 2.

Soient X, Y et Z trois variables aléatoires. Montrer que $X \to Y \to Z$ est une chaîne de Markov si et seulement si X et Z sont indépendantes conditionellement à Y.

Correction.

Par définition, cette suite de trois variables aléatoires est une chaîne de Markov si et seulement si elles vérifient

$$\mathbb{P}(Z = z|Y = y, X = x) = \mathbb{P}(Z = z|Y = y).$$

D'autre part, X et Z sont indépendantes conditionellement à Y si et seulement si elles vérifient

$$\mathbb{P}(X = x, Z = z|Y = y) = \mathbb{P}(X = x|Y = y)\mathbb{P}(Z = z|Y = y).$$

Or on peut voir que ces deux égalités sont équivalentes, grace à la formule de Bayes (utilisée 3 fois) (on utilise ici des notations allégées):

$$p(x,z|y) = \frac{p(x,y,z)}{p(y)} = \frac{p(x,y)p(z|x,y)}{p(y)} = \frac{p(y)p(x|y)p(z|x,y)}{p(y)} = p(x|y)p(z|x,y).$$

Exercice 3.
1) Déterminer si les matrices suivantes sont primitives (justifier votre réponse par les arguments de votre choix):

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1/3 & 0 & 2/3 \\ 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix}.$$

2) Soit $\mathbb{X} = (X_n)_{n\in\mathbb{N}}$ une châine de Markov homogène sur un ensemble fini, dont la matrice de transition est

$$M = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1/4 & 0 & 3/4 \end{pmatrix}.$$

Déterminer son taux d'entropie $H(\mathbb{X})$.

Correction.
1) Pour déterminer si ces matrices sont primitives, le plus simple est de dessiner le graphe associé et de vérifier s'il est irréductible (càd. peut-on se déplacer de tout sommet vers tout autre sommet?) et apériodique (cad. est-il vrai que le pgcd des longueurs des boucles centrées en tout point vaut 1?).
On voit assez facilement que tous ces graphes sont irréductibles (voir Figure 1).

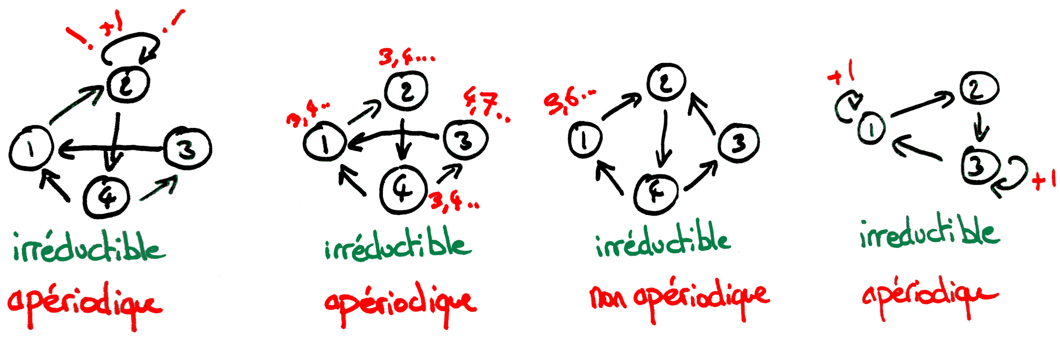

FIGURE 1. Graphes orientés correspondant aux matrices des questions 3.1 et 3.2.

Concernant l'apériodicité: Pour la première matrice, il y a une boucle en le sommet 2, ce qui assure que tout chemin peut être artificiellement rallongé de 1, donc elle est bien apériodique; Pour la deuxième matrice, aux sommets 1,2,4 il y a des boucles de longueur 3 et 4 (tels que pgcd(3,4)=1), et au sommet 3 il y a des boucles de longueur 4 et 7, donc elle est bien apériodique; Pour la troisième matrice, il y a un problème car depuis le sommet 2 on ne peut que faire des boucles de longueur multiple de 3, donc elle n'est pas primitive.
2) D'après le cours, pour connaitre le taux d'entropie il suffit de vérifier que M est primitive et de calculer sa loi stationnaire μ, ce qui donnera

$$H(X) = \sum_{i,j} \mu_i p_{ij} \log(p_{ij}) = \sum_i \mu_i H(\text{Ligne } i).$$

Ici il est facile de vérifier que le graphe de M est irréductible et apériodique (voir Figure 1), notamment grace aux boucles en 1 et 3. Ensuite on calcule μ en résolvant le système linéaire $\mu M = \mu$ sous la contrainte $\sum \mu_i = 1$, ce qui donne $\mu = (2/7, 1/7, 4/7)$. On en déduit

$$H_2(X) = \frac{2}{7}h(1/2) + \frac{1}{7}.0 + \frac{4}{7}h(1/4) = \frac{2}{7} + 0 + 2 - \frac{3}{4}\log(3) = \frac{16}{7} - \frac{3}{4}\log(3),$$

où on aura utilisé ici la notation $h(p) = -p\log(p) - (1-p)\log(1-p)$.