Master MFA première année 2011-2012

Module D4MA1C20, Compléments magistère
Feuille de TD n^{0}2

1. Codages uniquement décodables

Un codage $C : E \to \{0,1\}^*$ est dit *non singulier* si c'est une application injective. C'est une condition plus faible que d'être uniquement décodable. Les codages suivants de l'ensemble $E = \{1, 2, \dots, 8\}$ sont ils uniquement décodables ? Non singuliers ?

1–

x	1	2	3	4	5	6	7	8
$C(x)$	0	10	110	1110	1001	1010	1011	1100

2–

x	1	2	3	4	5	6	7	8
$C'(x)$	0	10	110	1110	111100	111110	111101	111111

3– Soit X une variable aléatoire à valeurs dans E, de distribution

x	1	2	3	4	5	6	7	8
$p(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$

Calculer son entropie. Calculer les longueurs moyennes des deux codages ci-dessus.

2. Codage de longueur moyenne inférieure à l'entropie

L'exercice précédent contient un exemple, sur un ensemble à 8 éléments, de variable aléatoire X à valeurs dans E et de codage $C : E \to \{0,1\}^*$ non singulier tels que $\mathbb{E}(\ell(X)) < H(X)$. Peut on faire plus simple ? Donner un exemple, le plus simple possible.

3. Codages non singuliers

Soit X une variable aléatoire à valeurs dans E. On s'intéresse à la borne inférieure $L_{ns}(X)$ des longueurs moyennes $\mathbb{E}(\ell(X))$ des codages non singuliers sur deux lettres.

1– Montrer que, pour tout codage non singulier, $\mathbb{E}(\ell(X)) \geqslant \frac{1}{\log_2 3} H(X) - 1$ (remarquer qu'en ajoutant une lettre ("espace"), tout codage non singulier devient uniquement décodable). Cette borne est elle atteinte ?

2– Montrer que le codage non singulier qui minimise $\mathbb{E}(\ell(X))$ est celui qui utilise au maximum les suites de lettres courtes, en attribuant les suites les plus courtes aux éléments de plus forte probabilité. En déduire la formule

$$L_{ns}(X) = \sum_{i=1}^{|E|} p(x_i) \lceil \log_2(\frac{i}{2} + 1) \rceil.$$

3– On note

$$M = \sum_{i=1}^{|E|} p(x_i) \log_2(\frac{i}{2} + 1).$$

Prouver l'encadrement $M \leqslant L_{ns}(X) < M+1$. Posant

$$c = \sum_{i=1}^{|E|} \frac{1}{\frac{i}{2}+1}, \quad q(x_i) = \frac{1}{c}\frac{1}{\frac{i}{2}+1},$$

utiliser $D(p||q)$ pour minorer $H(X) - M$.

4– Majorer c et en déduire l'encadrement

$$H(X) - \log_2 \log_2(|E|) - 2 \leqslant L_{ns}(X) < H(X) + 1.$$

4. Combien ai-je de jouets ?

Je cache un certain nombre de jouets dans mon tiroir. Nous jouons au jeu des 10 questions (je choisis un jouet, tu me poses des questions très astucieuses, je reponds par oui ou par non, jusqu'à ce que tu devines quel jouet j'ai choisi) de façon répétée. En moyenne, il te suffit de 6.65 questions. Peut on donner une minoration grossière du nombre de jouets que contient mon tiroir ?

5. Vin empoisonné

On doit décider laquelle parmi 6 bouteilles de vin contient un poison mortel. On peut dispose d'un certain nombre de rats de laboratoire, on peut leur faire goûter des mélanges. L'effet du poison est immédiat, même à faible dose. Un rat qui a goûté et survécu est saoûl donc ne peut être réutilisé. Il s'agit d'utiliser le moins de rats possible, sachant que la probabilité que le poison se trouve dans la i-ème bouteille est donnée par le tableau

i	1	2	3	4	5	6
$p(i)$	$\frac{8}{23}$	$\frac{6}{23}$	$\frac{4}{23}$	$\frac{2}{23}$	$\frac{2}{23}$	$\frac{1}{23}$

.

1– Quelle est l'espérance du nombre de rats utilisés ?

2– Par quel mélange faut il commencer ?

3– Quelle est l'espérance du nombre de rats sacrifiés ?

4– Le protocole retenu minimise t'il aussi l'espérance du nombre de rats sacrifiés ?

6. Jeu des 10 questions aléatoires

Je cache $m = 2^n$ jouets dans mon tiroir E, n grand. Nous jouons au jeu des 10 questions, mais tu es très bête, tu tires tes questions au hasard. Une question étant une fonction $f : E \to \{0,1\}$, la tirer au hasard signifie tirer indépendamment au hasard les valeurs $f(x)$, $x \in E$, suivant $\mathcal{B}(\frac{1}{2})$. On pose successivement k questions tirées indépendamment au hasard, et on obtient k fonctions $f_k : E \to \{0,1\}$.

1– Soit $x_0 \in E$. On note N le nombre de $x \in E \setminus \{x_0\}$ tels que pour tout $i = 1, \ldots, k$, $f_i(x) = f_i(x_0)$. Quelle est la loi de N ?

2– On suppose que $k = n-1$. Donner une approximation de la loi de N. En déduire que pour tout $x \in E$, avec probabilité bornée inférieurement, le jeu se termine en moins de $n-1$ questions.

3– On suppose que $k = n-1$. Montrer qu'il existe $x \in E$ pour lequel le jeu nécessite au moins n questions.