Théorie de l'information et du codage

TD n^{o}2 – Compression de données

Dans tout le sujet, $\mathcal{X}$ est un ensemble fini.

Exercice 1 – Divergence de Kullback-Leibler

On note $\mathcal{P}(\mathcal{X})$ l'ensemble des mesures de probabilité sur $\mathcal{X}$. On rappelle que la *distance en variation totale* entre deux éléments P et Q de $\mathcal{P}(\mathcal{X})$ est

$$||P - Q||_{\text{VT}} = \sup_{A \subseteq \mathcal{X}} |Q(A) - P(A)|.$$

1. Vérifier que pour tout $P, Q \in \mathcal{P}(\mathcal{X})$, l'on a :
$$||P - Q||_{\text{VT}} = \frac{1}{2} \sum_{x \in \mathcal{X}} |Q(x) - P(x)|.$$
2. La *divergence de Kullback-Leibler* de P par rapport à Q est définie comme :
$$D(P\|Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}.$$
Montrer que $D(P\|Q) \geq 0$ avec égalité si et seulement si $P = Q$. Est-ce une distance ?
3. Si P' et Q' désignent les lois images de P et Q par une fonction arbitraire $f \colon \mathcal{X} \to \mathcal{X}'$, comment se comparent $||P'-Q'||_{\text{VT}}$ et $||P-Q||_{\text{VT}}$? Qu'en est-il de $D(P'||Q')$ et $D(P||Q)$?
4. Établir l'inégalité de Pinsker, valable pour tout $P, Q \in \mathcal{P}(\mathcal{X})$:
$$D(P\|Q) \geq \frac{1}{2}||P - Q||^2_{\text{TV}}.$$
5. Exprimer l'information mutuelle $I(X, Y)$ de deux variables aléatoires discrètes X et Y comme une divergence de Kullback-Leibler.
6. Les codages étudiés jusqu'ici requièrent la connaissance exacte de la loi P de la source X, ce qui est irréalisable en pratique : on ne dispose que d'une certaine estimation à priori, que nous noterons Q. Calculer la perte de performance induite dans le cas où Q est dyadique, puis en donner un encadrement dans le cas où Q est quelconque.

Exercice 2 – Second principe thermodynamique

Soient $(X_n)_{n\geq 0}$ et $(Y_n)_{n\geq 0}$ deux chaînes de Markov de même matrice de transition sur $\mathcal{X}$.

1. Montrer que $D(\mathcal{L}(X_n)||\mathcal{L}(Y_n))$ décroît avec n, puis justifier le titre de l'exercice.
2. Lorsque la matrice de transition est irréductible et apériodique, retrouver le résultat bien connu que $\mathcal{L}(X_n)$ converge vers une loi stationnaire.

Exercice 3 – Débit entropique d'une source stationnaire

Soit $X = (X_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{X}$. On suppose que X est stationnaire au sens où pour tout entier $k \geq 1$ fixé, la loi du vecteur $(X_{n+1}, \ldots, X_{n+k})$ ne dépend pas de $n \geq 0$.

1. Montrer que la limite
$$H_\infty(X) = \lim_{n\to\infty} \frac{H(X_1, \ldots, X_n)}{n}$$
existe. On l'appelle le débit entropique de la source X.
2. Quelle signification a cette grandeur pour la compression de la suite X ?

Exercice 4 – Codage de Shannon

Shannon a proposé la stratégie suivante pour coder une source N-aire X de loi $P = (P_1, \ldots, P_N)$. On commence par réordonner les N valeurs possibles de X de telles sorte que :

$$P_1 \geq P_2 \geq \ldots \geq P_N > 0.$$

On définit alors la "fonction de répartition" de X comme le vecteur $(F_1, \ldots, F_N) \in [0,1)^N$, où pour tout $n \in \{1, \ldots, N\}$,

$$F_n = \sum_{i=1}^{n-1} P_i.$$

Enfin, on code chaque symbole source $n \in \{1, \ldots, N\}$ par l'écriture binaire du nombre F_n tronquée à $l_n = \lceil \log_2 \frac{1}{P_n} \rceil$ bits.

1. Montrer que ce code est instantané et que la longueur moyenne des mots-code vérifie :
$$H(X) \leq \mathbb{E}[L] < H(X) + 1.$$
2. Expliciter le code obtenu pour $P = (0.43, 0.22, 0.17, 0.09, 0.05, 0.04)$. Cette méthode est-elle optimale ?

Exercice 5 – Alphabet-source infini

On s'intéresse aux codes D-aires instantanés sur un alphabet infini $\mathcal{X} = \{1, 2, \ldots\}$. Étant donné un tel code, on note $l_1, l_2, \ldots$ les longueurs respectives des mots-code.

1. Montrer que les séquences d'entiers $l_1, l_2, \ldots$ possibles sont exactement celles qui satisfont l'inégalité de Kraft étendue :
$$\sum_{n=1}^{\infty} D^{-l_n} \leq 1.$$
2. En déduire que pour toute variable aléatoire X à valeurs dans $\mathcal{X}$, la longueur moyenne de n'importe quel code instantané de la source X vérifie :
$$\mathbb{E}[L] \geq H_D(X),$$
avec égalité si et seulement si $l_n = \log_D \mathbb{P}(X = n)$ pour tout $n \geq 1$.