M1 - THÉORIE DE L'INFORMATION - 2018-2019

TRAVAUX DIRIGÉS 1

Exercice 1. *(Fonction d'une variable aléatoire I)*

Geneviève se rend régulièrement à l'hippodrome, où se déroule une course entre huit chevaux. À la fin de la course, les gains de Geneviève dépendent de la performance du cheval sur lequel elle a misé: selon qu'il arrive en position $1, 2, 3, .., 8$, ses gains sont respectivement de $100, 50, 50, 0, -10$, $-50, -50, -100$ euros. On supposera toujours que tous les chevaux sont aussi bons les uns que les autres.

(1) On note X la variable aléatoire correspondant aux gains de Geneviève. Calculer l'espérance de ses gains $\mathbb{E}(X)$, ainsi que leur entropie $H_2(X)$.

Lassée de perdre de l'argent sur le long terme, Geneviève décide de monétiser son activité en mettant en ligne des vidéos de réaction aux résultats des courses. Elle passe un contrat avec un network: ses gains sont désormais sécurisés et correspondent à *la valeur absolue* de ses gains initiaux (car le public l'apprécie tout autant quand elle s'énerve après avoir perdu).

(2) On note $Y = |X|$ la variable aléatoire correspondant aux nouveaux gains de Geneviève. Calculer la fonction de masse q correspondant à ces gains.
On rappelle qu'étant donné une fonction f, la fonction de masse d'une variable aléatoire image $f(X)$ est donnée par

$$\mathbb{P}(f(X) = y) = \sum_{\{x \in \mathcal{X} \ | \ f(x)=y\}} \mathbb{P}(X = x).$$

(3) Calculer $\mathbb{E}(Y)$ et $H_2(Y)$. Que constate-t-on?
(4) Après chaque course, le montant des gains est transmis directement à la banque sous forme de code binaire. Les gains $(100, 50, 10, 0)$ sont codés en les mots $(00, 01, 10, 11)$. Quelle est la longueur moyenne du mot transmis?
(5) Voici un nouveau code: $(10, 0, 110, 111)$. Quelle est la longueur moyenne du mot transmis? Comparer avec le résultat de la question (3).
(6) *(Optionnel)* Essayer de trouver un autre code qui utilise une plus courte longueur moyenne.[1]

Exercice 2. *(Loi uniforme sur un ensemble fini)*

Soit X la variable aléatoire réelle qui suit une loi de probabilité uniforme sur un ensemble $\mathcal{X} = \{x_1, ..., x_n\}$ fini. Soit Y une autre variable aléatoire réelle sur $\mathcal{X}$, quelconque.

(1) Donner, pour tout $x \in \mathcal{X}$, la valeur de $p(x) = \mathbb{P}(X = x)$. En déduire la valeur de l'entropie $H(X)$.
(2) Soit $q : \mathcal{X} \to [0, 1]$ la fonction de masse de Y. Montrer que $D(q\|p) = H(X) - H(Y)$.
(3) Montrer que l'entropie que peut avoir une variable aléatoire sur $\mathcal{X}$ est toujours inférieure ou égale à $\log n$.
(4) Montrer que la loi uniforme est l'unique loi de probabilité maximisant cette entropie sur $\mathcal{X}$.

---

[1] On pourra jeter un oeil à `https://fr.wikipedia.org/wiki/Théorème_du_codage_de_source`.

Exercice 3. *(Fonction d'une variable aléatoire II)*

Soit X une variable aléatoire réelle qui suit une loi de probabilité sur un ensemble $\mathcal{X} = \{x_1, ..., x_n, ...\}$ discret. On notera p sa fonction de masse. Soit $\mathcal{Y}$ un autre ensemble discret, $f : \mathcal{X} \to \mathcal{Y}$ une fonction, et $Y := f(X)$ la variable aléatoire image de X par f.

(1) Montrer que $H(f(X)|X) = 0$, en utilisant la définition d'entropie conditionnelle.
(2) Montrer que $H(X, f(X)) = H(X)$.
(3) En déduire que $H(X) \geq H(f(X))$.
(4) Montrer que $H(X|f(X)) = 0$ si et seulement si f est injective.
(5) En déduire que $H(X) = H(f(X))$ si et seulement si f est injective.
(6) Supposons que Z soit une variable aléatoire sur $\mathcal{Y}$ telle que $H(Z|X) = 0$. Montrer qu'il doit exister une fonction $g : \mathcal{X} \to \mathcal{Y}$ telle que $Z = g(X)$.
Note: L'existence d'une telle fonction est équivalente à ce que pour tout $x \in \mathcal{X}$ tel que $p(x) > 0$, il existe un unique $y \in \mathcal{Y}$ vérifiant $\rho(x, y) > 0$; ici ρ désigne la fonction de masse de la loi conjointe (X, Y).

Exercice 4. *(Loi géométrique sur $\mathbb{N}$)*

On tire une pièce à pile ou face pendant plusieurs tours, et ce jusqu'à ce que l'on obtienne une face. On note X la variable aléatoire sur $\mathbb{N}$ qui désigne le numéro du tour au cours duquel on a obtenu une face. Pour des raisons pratiques, on considère que le premier tirage correspond au tour 0, le second tirage au tour 1, etc. On suppose que la pièce est déséquilibrée: la probabilité de tirer une face à chaque tour est de $\alpha \in]0, 1[$.

(1) Calculer $\mathbb{P}(X = 0), \mathbb{P}(X = 1), \mathbb{P}(X = 2)$. Calculer, pour tout $k \in \mathbb{N}$, $\mathbb{P}(X = k)$.
(2) Calculer l'espérance $\mathbb{E}(X)$. *(On pourra utiliser la relation $\sum_{k \in \mathbb{N}} k\beta^k = \frac{\beta}{(1-\beta)^2}$ pour $\beta < 1$.)*
(3) Montrer que $H(X) = -\log(\alpha) - \frac{1-\alpha}{\alpha} \log(1 - \alpha)$.
(4) Soit $h(\alpha)$ l'entropie de X lorsque la probabilité de tirer face est α. Montrer que h est décroissante sur $]0, 1[$.
(5) Calculer la limite de h lorsque α tend vers 0 et 1. Existe-t-il une loi d'entropie maximale sur $\mathbb{N}$? *(On pourra utiliser le fait que $\lim_{t \to 0} t \log t = 0$)*
(6) Soit Y une variable aléatoire sur $\mathbb{N}$, telle que $\mathbb{E}(Y) = m > 0$. On note p et q les fonctions de masse de X et Y, respectivement. Montrer que $H(X) - H(Y) = D(q\|p)$. *(comparer avec l'exercice 2)*
(7) Montrer que parmi toutes les lois sur $\mathbb{N}$ ayant pour espérance $m > 0$, la loi géométrique X avec $\alpha = \frac{1}{1+m}$ est l'unique loi maximisant l'entropie. Exprimer cette entropie maximale en fonction de m.

M1 - THÉORIE DE L'INFORMATION - 2018-2019

TRAVAUX DIRIGÉS 1 (SUITE)

Exercice 5. *(Mélanger accroit l'entropie)*
On considère X une v.a. sur un jeu de cartes $\mathcal{X} = \{1, \dots, 52\}$. On considère également $\mathcal{S}_{52}$ le groupe des permutations sur $\mathcal{X}$, et S une variable aléatoire sur $\mathcal{S}_{52}$. On note $Y = SX$, la variable désignant un jeu de cartes qui à subi un mélange aléatoire. Montrer que $H(Y) \geq H(X)$.

Exercice 6.
Une urne contient b boules blanches et r boules rouges. On note X (resp. Y) la variable aléatoire correspondant au tirage de 5 boules dans l'urne avec replacement (resp. sans replacement). Laquelle de ces deux variables a la plus grande entropie? Justifier.

Exercice 7. *(Piocher et tirer I)*
On considère un sac contenant deux pièces lestées, que nous appellerons a et b. Chacune a une probabilité de faire face égale à p et q, respectivement ($p, q \in]0, 1[$). On choisit uniformément au hasard une pièce dans le sac, et on la tire à pile ou face plusieurs fois. On note Y la v.a. sur $\{a, b\}$ désignant la pièce qui a été choisie, et X_n la v.a. sur $\{\text{pile, face}\}$ désignant le résultat du n-ème tirage.

(1) Montrer que X_1 et X_2 sont identiquement distribuées, en calculant leur lois de probabilité.
(2) X_1 et X_2 sont-elles indépendantes? On donnera la réponse en fonction de p et q.
(3) Que vaut $I(X_1; X_2|Y)$?

Exercice 8. *(Somme de variables aléatoires)*
Soit X une v.a. sur $\mathcal{X} = \{x_1, ..., x_m\} \subset \mathbb{R}$ et Y une v.a. sur $\mathcal{Y} = \{y_1, ..., y_n\}$. Soit $Z = X + Y$.

(1) Vérifier que $H(Z) \leqslant H(X) + H(Y)$.
(2) Montrer que $H(Z|X) = H(Y|X)$.
(3) Si X et Y sont indépendantes, en déduire que $\max\{H(X); H(Y)\} \leqslant H(Z)$.
(4) Trouver un exemple pour lequel $\min\{H(X); H(Y)\} > H(Z)$.

Exercice 9.
On rappelle que $\rho : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ est une métrique sur $\mathcal{X}$ ssi pour tout $(x, y, z) \in \mathcal{X}^3$:

i) $\rho(x, y) \geq 0$,
ii) $\rho(x, y) = \rho(y, x)$,
iii) $\rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$,
iv) $\rho(x, y) = 0 \Leftrightarrow x = y$.

Étant données X et Y deux v.a. discrètes sur $\mathcal{X}$, on définit $\rho(X, Y) := H(X|Y) + H(Y|X)$.

(1) Montrer que ρ vérifie les propriétés i),ii) et iii).
(2) Trouver une condition nécessaire et suffisante sur X et Y pour que $\rho(X, Y) = 0$. Est-ce une métrique?

Exercice 10.
Étant données deux v.a. X et Y identiquement distribuées, on pose $\rho(X; Y) = 1 - \frac{H(Y|X)}{H(X)}$.

(1) Montrer que $\rho(X; Y) = \frac{I(X;Y)}{H(X)}$. En déduire que ρ est symétrique.
(2) Montrer que $\rho(X; Y) \in [0, 1]$. A quoi correspondent les cas $\rho(X; Y) = 0, 1$?

Exercice 11.
Soit $\mathbf{X} = (X_n)_{n\in\mathbb{N}}$ un processus stochastique sur un ensemble $\mathcal{X}$ fini (de taille $k \geq 1$). Montrer que le taux d'entropie supérieur $\overline{H}(\mathbf{X})$ est toujours inférieur ou égal à $\log k$. Cette borne est-elle optimale? Que se passe-t-il si on remplace $\mathcal{X}$ par $\mathbb{N}$?

Exercice 12. *(Chaine de Markov homogène à deux états I)*
Soit $\mathbf{X} = (X_n)_{n\in\mathbb{N}}$ une chaine de Markov homogène sur $\mathcal{X} = \{a, b\}$.
On note $p = \mathbb{P}(X_1 = a|X_0 = b)$ et $q = \mathbb{P}(X_1 = b|X_0 = a)$.

(1) Écrire P, la matrice de transition de cette chaine de Markov, et tracer son graphe.
(2) Montrer que l'unique distribution stationnaire de $H(\mathbf{X})$ est $\bar{\mu} = (\frac{p}{p+q}, \frac{q}{p+q})$.
(3) *Pour le reste de cet exercice, on suppose que X_0 suit la distribution μ.*
Exprimer, pour tout $n \in \mathbb{N}$, l'entropie $H(X_n)$ en fonction de p et q. On pourra exprimer le résultat avec fonction $h(t) = -t\log(t) - (1-t)\log(1-t)$.
(4) Calculer le taux d'entropie $H(\mathbf{X})$.
(5) Quelles valeurs pour p et q maximisent le taux d'entropie?

Exercice 13. *(Piocher et tirer II)*
On reprend les notations et le contexte de l'exercice 7. On note $\mathbf{X} = (X_n)_{n\in\mathbb{N}^*}$ le processus aléatoire correspondant au tirage de la pièce que l'on a piochée.

(1) Calculer $\lim\limits_{n\to+\infty} \frac{1}{n} H(X_1, \ldots, X_n|Y)$.
(2) En déduire le taux d'entropie $H(\mathbf{X})$.
(3) Est-ce que $\mathbf{X}$ est une chaine de Markov homogène?

Exercice 14. *(Chaine de Markov homogène à deux états II)*
On reprend le même contexte que l'exercice précédent. On note $\mu_n = (u_n, v_n)$ la distribution que suit la variable X_n. Ici, on ne suppose pas que μ_0 est stationnaire, et on s'intéresse à la convergence de la suite μ_n.

(1) Exprimer μ_n en fonction de P et μ_0.
(2) $(\mu_n)_{n\in\mathbb{N}}$ admet-elle une limite lorsque $(p, q) = (1, 1)$ ou $(0, 0)$? Justifier, et calculer cette limite le cas échéant.
(3) *On suppose à partir de maintenant que $p + q \in]0, 2[$.*
Calculer le spectre de P.
(4) Trouver une base de $\mathbb{R}^2$ composée de vecteurs propres de P, puis calculer P^n.
(5) En déduire la limite de μ_n lorsque $n \to +\infty$. Que constate-t-on?