TD 1 : Entropie et Infromation

Yann Rotella

2026

Cette feuille d'exercices est prévue pour deux séances. Certains exercices sont des redites du cours. A priori, ces exercices sont réalisables sans le poly du cours, et ce qui est demandé provenant du cours doit être appris sans avoir besoin de revoir le cours.

Première partie

Exercice 1. *Une application simple.*
On considère un dé à 6 faces équilibré.

1. Donner l'information mutuelle moyenne entre la face du haut et la face cachée du dé.
2. Donner l'information mutuelle moyenne entre la face du haut du dé et la face tournée vers le joueur.

Exercice 2. *Dualité information et incertitude.*
On va maintenant trouver théoriquement la relation entre l'entropie, l'entropie conditionnelle et l'information mutuelle moyenne.

1. Rappeler les formules de $I(X;Y)$, $H(X)$ et $H(X|Y)$.
2. Donner la relation entre ces trois quantités et montrer la.

Exercice 3. *Jeu de cartes.*
Un jeu de 32 cartes comporte :

- 8 cartes de coeur ♡,
- 8 cartes de carreau ♢,
- 8 cartes de pique ♠,
- et 8 cartes de trèfle ♣.

Ces 8 cartes sont, par valeur décroissante, l'as, le roi, la dame, le valet, le 10, le 9, le 8 et le 7. On considère une "main" de 4 cartes tirées au hasard d'un jeu de 32 cartes, ainsi que les évènements suivants.

— E_1 : la main ne contient aucune carte inférieure au valet,
— E_2 : le main ne contient pas de figure (roi, dame, valet),
— E_3 : la main contient 4 cartes du même nom.
— E_4 : la main contient les 4 as.

1. Calculer la quantité d'information propre $I(E_i)$ associée à chaque évènement E_i, $i \in \{1,2,3,4\}$.
2. Calculer les informations mutuelles $I(E_1, E_2)$ et $I(E_1, E_3)$.
3. Évaluer approximativement la quantité d'information nécessaire pour spécifier une main de 4 cartes. Comparer cette quantité à l'entropie de la variable aléatoire correspondant au contenu d'une main.

Exercice 4. *Règle de chaînage de l'entropie.*
Soit un vecteur $(X_1, X_2 \ldots, X_n)$ de n variables aléatoires. Par définition on sait que son entropie est

$$H(X_1, X_2 \ldots, X_n) = -\sum p(x_1, \ldots, x_n) \log(p(x_1, \ldots, x_n)) .$$

1. (Cas d'indépendance) Montrer que si les variables $X_1, X_2 \ldots, X_n$ sont indépendantes, alors :

$$H(X_1, X_2 \ldots, X_n) = \sum_{i=1}^{n} H(X_i).$$

2. (Cas général : « règle de chaînage pour l'entropie ») Montrer que :

$$H(X_1, X_2 \ldots, X_n) = H(X_n|X_1, \ldots, X_{n-1}) + H(X_{n-1}|X_1, \ldots, X_{n-2}) + \cdots + H(X_2|X_1) + H(X_1).$$

Exercice 5.
Soit X une variable aléatoire et g une fonction.

1. En utilisant le règle de chaînage de deux manières différentes, montrer que

$$H(g(X)) \leq H(X).$$

2. Dans quelle condition a-t-on l'égalité ?

Exercices complémentaires (partie 1), à la maison

Exercice 6. *Intuition de l'information propre.*
Expliquer en détails pourquoi l'information de l'événement $X = x$ est donnée par la formule

$$I(x) = \log_b(\frac{1}{p_X(x)})$$

où $x \in \mathcal{X}$ et $(\mathcal{X}, p_X)$ est un espace probabilisé.

Exercice 7. *Probabilités et Bernouilli.*
On considère n variables aléatoires $X_1, \ldots, X_n$ dans $\{0, 1\}$ suivant une loi de Bernouilli de paramètre p. Donner l'espérance de la somme de ces variables aléatoires.

Exercice 8. *Questions de cours.*
1. Donner la définition d'un bit d'information.
2. Soient X et Y deux v.a. Donner la définition mathématique de l'information mutuelle moyenne, de l'entropie de X sachant Y et de l'entropie de X, puis donner et montrer la relation entre ces trois quantités.
3. Montrer que pour toute v.a X dans $\mathcal{X}$ où $|\mathcal{X}| = k$, on a

$$H(X) \leq \log_2(k)\,.$$

Exercice 9. *Application numérique à trois variables.*
Soient X, Y, Z trois v.a. dans $\{0, 1\}$ avec

$$p_{XYZ}(0, 0, 1) = \frac{1}{4} = p_{XYZ}(1, 0, 1) = p_{XYZ}(0, 1, 1) = p_{XYZ}(1, 0, 0)$$

Calculez $H(X)$, $H(Y|X)$, $H(Z|X, Y)$. En déduire $H(X, Y, Z)$. Calculez $H(Y)$. Combien d'information apporte X sur Y et réciproquement ?

Exercice 10. *Un problème de météo.*
Dans la vallée de la mort :
— il pleut en moyenne 1 jour sur 100.
— la météo prédit 3 jours de pluie sur 100.
— chaque fois qu'il pleut, la météo l'a prévu.
Monsieur Sûr-de-lui prévoit qu'il ne pleut jamais. Est-il justifié de payer cher des investissements météo, alors que Monsieur Sûr-de-lui, qui ne coûte rien et se trompe moins souvent que la météo ?

Deuxième partie - exercices plus complets

Exercice 11. *Urnes.*
On a $n+1$ urnes contenant chacune $0, 1, \ldots, n$ boules blanches et $n, n-1, \ldots, 0$ boules noires. On choisit une urne au hasard. On tire ensuite une boule blanche.

1. Quelle est la probabilité que l'urne tirée soit l'urne i ?
2. Quelle est la probabilité que la prochaine boule tirée soit blanche (en ayant remis la boule et en ayant choisi la même urne) ?

Exercice 12. *Pesées.*
On considère un ensemble de n pièces d'or. Parmis ces pièces une seule est fausse et a un poids inférieur au poids standard. De plus on dispose d'une balance à deux plateaux permettant de comparer les poids a et b de deux ensembles A et B de pièces posés respectivement sur chacun des plateaux.

1. Quelle est la quantité d'information nécessaire pour déterminer la fausse pièce ?
2. On suppose dans cette question que $n = 3k$. Calculer la quantité d'information qu'apporte une pesée quand $|A| = |B| = k$.
3. En déduire une borne inférieure du nombre moyen m de pesées nécessaires pour déterminer la fausse pièce. Que peut-on dire si n est de la forme 3^i ?

Exercice 13. *Entropie et somme de variables.*
Soient X et Y deux variables aléatoires à valeurs dans un groupe $(G, +)$. Soit la variable aléatoire $Z = X + Y$.

1. Montrer que $H(Z|X) = H(Y|X)$.
2. Montrer que si X et Y sont indépendantes alors $H(Y) \leq H(Z)$ et $H(X) \leq H(Z)$ (utiliser la positivité de l'information mutuelle).
3. Donner un exemple de deux variables aléatoires X et Y telles que $H(X) > H(Z)$ et $H(Y) > H(Z)$.

Exercice 14. *Le problème du mot de passe.*
Un individu (probablement mal intentionné) cherche à accéder à un service protégé par un mot de passe qu'il ne connaît pas. Soit $\mathcal{M} = \{0,1\}^m$ l'ensemble des mots de passe possibles. Nous supposons que le système d'authentification est parfait et que la seule possibilité d'action pour l'attaquant consiste à essayer les mots de passe un par un.

On suppose ensuite que le mot de passe est choisi dans $\mathcal{M}$ selon une loi d'entropie $h \leq m$. Nous notons p_i les probabilités des mots de $\mathcal{M}$dans l'ordre décroissant (le mot le plus probable a pour probabilité p_1, le suivant p_2 ...).

1. Montrer que la meilleure stratégie consiste à tester les mots dans l'ordre des probabilités décroissantes. Exprimez le nombre moyen d'essais, $\mathcal{N}(p)$, en fonction des p_i.
2. Soient p et q deux distributions de probabilités, on note
$$D(p||q) = \sum_x p(x) \log_2 \left(\frac{p(x)}{q(x)} \right)$$
la divergence de Kullback-Leibler. Montrer que $D(p||q) \geq 0$ avec égalité si et seulement si $p = q$.
3. Soient deux lois de probabilité $p = (p_i)_{i \geq 1}$ et $q = (q_i)_{i \geq 1}$ telles que les suites p_i et q_i soient décroissantes avec $q_i > 0$ pour tout $i \geq 1$ (en revanche p_i peut être nul à partir d'un certain rang). Nous posons $q_i = (1-\alpha)\alpha^{i-1}$ pour un certain réel $0 < \alpha < 1$. On suppose que les entropies $H(p)$ et $H(q)$ sont bien définies. Montrer que si $H(p) = H(q)$ alors
$$\sum_{i \geq 1} i p_i \geq \sum_{i \geq 1} i q_i.$$

4. Calculer l'entropie $H(q)$ de la loi q en fonction de α. Nous noterons H_α cette quantité. On rappelle les identités $\sum_{i\geq 1} \alpha^{i-1} = \frac{1}{1-\alpha}$ et $\sum_{i\geq 1} \alpha^{i-1} = \frac{1}{(1-\alpha)^2}$.
5. En déduire que pour tout réel $0 < \alpha < 1$ nous avons $1 < (1-\alpha)2^{H_\alpha} < e$, où e est la base du logarithme népérien.
6. Déduire du résultat précédent que $\mathcal{N}(p) \geq c_1 2^h$ (on s'efforcera de donner une valeur à c_1).
7. Expliquer en français l'implication de ce résultat.

Exercice 15. *Théorème de Shannon.*
Un chiffrement est incondtionnellement sûr si la connaissance d'un chiffré n'apporte **aucune** information sur le message clair. Le théorème de Shannon (en cryptographie) peut s'exprimer comme suit.

Soit $\mathcal{M}$, $\mathcal{C}$ et $\mathcal{K}$ trois espaces finis de même taille N décrivant respectivement l'espace des messages clairs, l'espace des messages chiffrés et l'espace des clefs secrètes. Alors un chiffrement e est incondtionnellement sûr si et seulement si

(i) K la v.a. des clefs secrètes suit une loi uniforme.
(ii) Pour tout $(m, c) \in \mathcal{M} \times \mathcal{C}$, il existe une unique clef $k \in \mathcal{K}$ telle que $E_k(m) = c$.

1. Exprimer la condition incondtionnellement sûr avec la théorie de l'information.
2. Montrer le théorème de Shannon.
3. Montrer que si l'espace des messages est plus grand que l'espace des clefs, alors aucun système n'admet une sécurité parfaite.
4. Est-ce que K et M (la v.a. correspondante aux messages clairs) sont indépendantes en pratique ?
5. Montrer que
$$H(K|M, C) \geq H(K) - H(C)$$
6. En supposant M et C indépendantes, que vaut $H(K|M, C)$? Quelle est l'interprétation de cette valeur ? Avec $\{0,1\}^n = \mathcal{M} = \mathcal{C}$, et une loi uniforme le tout et $\mathcal{K} = \{0,1\}^\kappa$.

Exercices complémentaires

Exercice 16. *Entropie et tirages uniformes.*
Soit X une variable aléatoire de distribution $(p_1, p_2, p_3, p_4) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$.

1. Quelle est l'entropie de X ? Quelle serait l'entropie maximale d'une variable aléatoire à image dans un ensemble à 4 éléments ?

   Soit U la variable aléatoire uniforme sur $\{0,1\}$. On considère l'algorithme suivant :
   (i) Tirer $u_1 \leftarrow U$. Si $u_1 = 0$, retourner $x = 1$.
   (ii) Sinon, tirer $u_2 \leftarrow U$. Si $u_2 = 0$, retourner $x = 2$.
   (iii) Sinon tirer $u_3 \leftarrow U$. Si $u_3 = 0$, retourner $x = 3$. Sinon, retourner $x = 4$.
   On suppose que les tirages successifs sont indépendants.
2. Démontrer que l'agorithme retourne la valeur i avec probabilité p_i.
3. Quel est le nombre moyen de tirages dans U lors d'une exécution de l'algorithme ?

Exercice 17. *Asymptotique de coefficients binomiaux.*
Pour $\lambda > 0$, on note $h(\lambda) = -\lambda \log_2(\lambda) - (1-\lambda)\log_2(1-\lambda)$. Dans cet exercice, on s'intéresse à l'asymptotique des coefficients de la forme
$$\binom{\lambda n}{n}$$
1. Démontrer que pour tout $n \geq 0$,
$$\sum_{i \leq \lambda n} \binom{i}{n} \leq 2^{nh(\lambda)}$$
   puis en déduire une borne supérieure sur $\binom{\lambda n}{n}$.

2. En utilisant la formule de Stirling, donner un équivalent asymptotique de $\frac{1}{n}\log_2\binom{\lambda n}{n}$.

Exercice 18. *Perte et traitement de l'information.*
On peut dire, que lorsque on a une perte d'information d'une source, celle-ci est perdue pour toujours. L'idée de cet exercice est de le montrer.

Une chaîne de Markov d'ordre 1 est une suite de variables aléatoires telles que

$$\forall n \in \mathbb{N}, P(X_{n+1} = x_n | X_n = x_n, \ldots, X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = x_n)$$

Soit $X \to Y \to Z$ une chaîne de Markov.

1. Montrer que $Z \to Y \to X$ est une chaîne de Markov.
2. Montrer que
$$H(X) \geq I(X,Y) \geq I(X,Z)\,.$$
Interpréter ce résultat.
3. Soit $g : \mathcal{X} \to \mathcal{X}$. Montrer que
$$I(X,Y) \geq I(X,g(Y))\,.$$
Interpréter ce résultat.
4. Imaginer un contexte actuel où on a une perte d'information de traitement.