DM 3

Exercice 1 (Maximum d'entropie)**.** Soit X une variable aléatoire réelle de densité f_X. Son entropie de Boltzmann-Shannon est définie par

$$H(X) := -\int f_X(x)\ln(f_X(x))\,dx$$

(avec la convention $0\ln(0) = 0$) lorsque la fonction $f_X \ln(f_X)$ est intégrable. On dit dans ce cas que X est d'entropie finie.

1. Montrer que si X est d'entropie finie alors pour tout $a \neq 0$ et $b \in \mathbb{R}$, $H(aX+b) = H(X)+\ln(|a|)$.
2. Calculer $H(X)$ dans chacun des cas suivants :
   (a) $X_1 \sim \mathcal{N}(m,\sigma^2)$, $m \in \mathbb{R}$, $\sigma > 0$;
   (b) $X_2 \sim \mathcal{U}([a,b])$, $a < b$;
   (c) $X_3 \sim \mathcal{E}(\lambda)$, $\lambda > 0$.
3. On considère une variable aléatoire X admettant une densité de la forme $f_X(x) = e^{-V(x)}\mathbb{1}_I(x)$ avec I un intervalle de $\mathbb{R}$ et $V : I \to \mathbb{R}$ une fonction mesurable telle que $\int_I e^{-V(x)}\,\mathrm{d}x = 1$.
   (a) Soit Y une variable aléatoire à valeurs dans I et d'entropie finie. On suppose que Y admet une densité notée f_Y telle que $V(Y)$ est intégrable et vérifie $\mathbb{E}(V(Y)) = \mathbb{E}(V(X))$. Montrer que
   $$H(X) - H(Y) = -\int_{\{f_Y>0\}} \ln(f_X(x)/f_Y(x))\,f_Y(x)\,dx.$$
   En appliquant l'inégalité de Jensen, en déduire que $H(Y) \leq H(X)$.
   (b) En déduire que X_1 est la variable aléatoire d'entropie maximale parmi les variables Y d'entropie finie telles que $\mathrm{Var}(Y) = \sigma^2$. Enoncer une propriété analogue pour les variables X_2 et X_3.

Solution.

1. Soit $a \neq 0$ et $b \in \mathbb{R}$. On commence par donner la densité de $aX+b$ en fonction de celle de X. On utilise la méthode de la fonction muette. Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction continue à support compact. On a
   $$\mathbb{E}g(aX+b) = \int g(ax+b)f_X(x)dx = \int g(u)f_X\left(\frac{u-b}{a}\right)\frac{du}{|a|}.$$
   On en déduit que $aX+b$ admet une densité qui est donnée par
   $$u \in \mathbb{R} \to \frac{1}{|a|}f_X\left(\frac{u-b}{a}\right).$$
   L'entropie de $aX+b$ est alors donnée par
   $$H(aX+b) = -\int f_{aX+b}(x)\ln f_{aX+b}(x)dx = -\int \left(\ln f_X\left(\frac{u-b}{a}\right) - \ln(|a|)\right) f_X\left(\frac{u-b}{a}\right)\frac{du}{|a|}.$$
   On en déduit que $H(aX+b) = H(aX) = H(X) + \ln(|a|)$.
2. Soit $X_1 \sim \mathcal{N}(m,\sigma^2)$. Soit $g \sim \mathcal{N}(0,1)$. On a $X_1 \sim \sigma g + m$. D'après la question précédente, on a $H(X_1) = H(\sigma g + m) = H(g) + \ln(\sigma)$. Il reste à calculer l'entropie de g. On a
   $$\begin{aligned}H(g) &= -\int \ln\left(\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right)\right)\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right)dx = \frac{1}{2}\ln(2\pi) + \int \frac{x^2}{2}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right)dx \\ &= \frac{1}{2}\ln(2\pi) + \frac{1}{2}.\end{aligned}$$

On obtient alors $H(X_1) = \frac{1}{2}\ln(2\pi e) + \ln(\sigma)$.

Soit $X_2 \sim \mathcal{U}([a,b])$. La v.a. X_2 admet une densité par rapport à la mesure de Lebesgue donnée par $f : x \to (b-a)^{-1} I(x \in [a,b])$. On a

$$H(X_2) = -\int \ln(f(x)) f(x) dx = -\int_a^b \ln\left(\frac{1}{b-a}\right) \frac{dx}{b-a} = \ln(b-a).$$

On a donc $H(X_2) = \ln(b-a)$.

Soit $X_3 \sim \mathcal{E}(\lambda)$ pour $\lambda > 0$. La v.a. X_3 admet une densité par rapport à la mesure de Lebesgue donnée par

$$f : x \in \mathbb{R} \to \lambda \exp(-\lambda x) I(x > 0).$$

L'entropie de X_3 est donnée par

$$\begin{aligned} H(X_3) &= -\int f(x)\ln(f(x))dx = -\int_0^{+\infty} \ln(\lambda\exp(-\lambda x))\lambda\exp(-\lambda x)dx \\ &= -\ln(\lambda) + \int_0^{+\infty} (\lambda x)\lambda \exp(-\lambda x)dx = 1 - \ln(\lambda). \end{aligned}$$

On a donc $H(X_3) = 1 - \ln(\lambda)$.

3.(a) Pour tout $x \in I$, $\ln(f_X(x)) = -V(x)$. On a donc

$$-H(X) = \mathbb{E}\ln(f_X(X)) = -\mathbb{E}V(X) = -\mathbb{E}V(Y) = \mathbb{E}\ln(f_X(Y)).$$

On obtient

$$\begin{aligned} H(X) - H(Y) &= -\mathbb{E}\ln(f_X(Y)) + \mathbb{E}\ln(f_Y(Y)) = -\mathbb{E}\ln(f_X(Y)/f_Y(Y)) \\ &= -\int_{\{f_Y>0\}} \ln(f_X(x)/f_Y(x))\, f_Y(x)\, dx. \end{aligned}$$

La fonction ln étant concave, l'inégalité de Jensen nous donne

$$\int_{\{f_Y>0\}} \ln(f_X(x)/f_Y(x))\, f_Y(x)\, dx \le \ln \int_{\{f_Y>0\}} f_X(x)\, dx \le \ln \int f_X(x)\, dx = 0.$$

D'où $H(Y) \le H(X)$. En d'autres termes, X est d'entropie maximale parmi toutes les v.a. Y à valeurs dans I, d'entropie finie et telles que $\mathbb{E}(V(Y)) = \mathbb{E}(V(X))$.

3.(b) Si Y est une variable aléatoire d'entropie finie telle que $\mathrm{Var}(Y) = \sigma^2$, alors d'après la question précédente avec $V(x) = \frac{x^2}{2\sigma^2} + \text{constante}$, $x \in I = \mathbb{R}$, on a $H(Y-m) \le H(X_1)$. Mais comme $H(Y-m) = H(Y)$, on a bien $H(Y) \le H(X_1)$.

De même, X_2 maximise l'entropie parmi toutes les variables aléatoires d'entropie finie à valeurs dans $[a,b]$ et X_3 parmi toutes les variables aléatoires d'entropie finie et intégrables à valeurs dans $\mathbb{R}^+$ dont l'espérance vaut $1/\lambda$.