DNS 2 : pour le vendredi 11 octobre

Le candidat encadrera ou soulignera les résultats.
N.B. : le candidat attachera la plus grande importance à la clarté, à la précision et à la concision de la rédaction. Si un candidat est amené à repérer ce qui peut lui sembler être une erreur d'énoncé, il le signalera sur sa copie et devra poursuivre sa composition en expliquant les raisons des initiatives qu'il a été amené à prendre.

Exercice 1 *(Problème : Entropie au sens de Shannon)*.

I. **Préliminaire**

I.A Représenter graphiquement la fonction logarithme népérien.

I.B Démontrer que, pour tout $x \in \mathbb{R}_+^\star$, $\ln(x) \leq x - 1$ et que $\ln(x) = x - 1$ si et seulement si $x = 1$.

I.C Donner une interprétation graphique de ces deux résultats.

I.D Montrer que la fonction g définie sur $[0,1]$ par $g(0) = 0$ et $\forall x \in]0,1]$, $g(x) = x\ln(x)$ est continue sur $[0,1]$ et dérivable sur $]0,1]$. Représenter graphiquement la fonction g.

On admet, pour tout $q \in]-1,1[$, que la série $\sum nq^{n-1}$ converge absolument et que $\displaystyle\sum_{k=1}^{+\infty} kq^{k-1} = \frac{1}{(1-q)^2}$.

II. **Entropie d'une variable aléatoire**

II.A Dans cette sous-partie, toutes les variables aléatoires considérées sont définies sur un même univers fini Ω et prennent leurs valeurs dans $[\![0,n]\!]$.
Si X est une telle variable, on note $p_k = \mathbb{P}(X = k)$. On définit l'entropie de X par :

$$H(X) = -\sum_{k=0}^{n} p_k \ln(p_k)$$

en convenant que $p_k \ln(p_k)$ vaut 0 lorsque $p_k = 0$.

II.A.1) Montrer que $H(X) \geq 0$ et que $H(X) = 0$ si et seulement si X est une variable aléatoire certaine, c'est-à-dire

$$\exists i \in [\![0,n]\!] \qquad \text{tel que} \qquad p_i = 1 \quad \text{et} \quad \forall j \neq i, \quad p_j = 0.$$

II.A.2) (a) X_0 est une variable aléatoire suivant la loi uniforme sur $[\![0,n]\!]$. Calculer $H(X_0)$.

(b) En appliquant l'inégalité de la question I.B à un nombre réel x bien choisi, démontrer que

$$\forall k \in [\![0,n]\!] \qquad -p_k\ln(p_k) + p_k \ln\left(\frac{1}{n+1}\right) \leq \frac{1}{n+1} - p_k.$$

(c) En déduire que $H(X) \leq H(X_0)$ avec égalité si et seulement si X suit la même loi que X_0 (pour le cas d'égalité on pourra utiliser le cas d'égalité de la question I.B).

II.B Dans cette sous-partie, on s'intéresse à des variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathbb{P})$ et prenant leurs valeurs dans $\mathbb{N}^\star$. Si X est une telle variable pour laquelle $\mathbb{P}(X = k)$ est noté p_k, on dit qu'elle est d'entropie finie si la série $\sum p_k \ln(p_k)$ est absolument convergente et on définit alors son entropie par

$$H(X) = -\sum_{k=1}^{+\infty} p_k \ln(p_k)$$

en convenant à nouveau que $p_k \ln(p_k)$ vaut 0 lorsque $p_k = 0$.

II.B.1) Pour $p \in]0,1[$, X_1 est une variable aléatoire suivant la loi géométrique de paramètre p.
Rappeler les valeurs de $\mathbb{P}(X_1 = k)$ et de l'espérance de X_1 (aucune démonstration n'est demandée).
Démontrer que X_1 est d'entropie finie et que $H(X_1) = -\dfrac{1-p}{p}\ln(1-p) - \ln(p)$.

II.B.2) Dans cette question et la suivante, X est une variable aléatoire à valeurs dans $\mathbb{N}^\star$ d'espérance finie.
On note $\mathbb{E}(X) = \displaystyle\sum_{k=1}^{+\infty} kp_k$. On se propose de démontrer que X est d'entropie finie.

(a) Quelle est la limite de p_k lorsque k tend vers $+\infty$?

(b) En déduire que $\displaystyle\lim_{k\to+\infty} \sqrt{p_k}\ln(p_k) = 0$, puis qu'il existe un entier k_0 tel que $\forall k \geq k_0 \quad 0 \leq -\sqrt{p_k}\ln(p_k) \leq 1$.

(c) Soit $k \geq k_0$. Montrer que

— si $p_k \leq \dfrac{1}{k^3}$, alors $0 \leq -p_k\ln(p_k) \leq \dfrac{1}{k^{3/2}}$;

— si $p_k \geq \dfrac{1}{k^3}$, alors $0 \leq -p_k \ln(p_k) \leq 3p_k \ln(k)$.

(d) Soit $k \geq 1$. Justifier que $\ln(k) \leq k$, puis que la série $\displaystyle\sum_{k\geq 1} \left(\frac{1}{k^{3/2}} + 3p_k \ln(k)\right)$ converge.

(e) Conclure.

II.B.3) Dans cette question, on suppose en plus que $\mathbb{E}(X) \leq 1/p$, p étant un réel de l'intervalle $]0,1[$. On veut montrer que $H(X) \leq H(X_1)$ (entropie d'une variable aléatoire suivant la loi géométrique de paramètre p dont la valeur a été calculée à la question II.B.1).
Pour $k \in \mathbb{N}^\star$, on note $p_k = \mathbb{P}(X = k)$ et $q_k = \mathbb{P}(X_1 = k)$.

(a) Justifier que la série $\displaystyle\sum_{k=1}^{+\infty}(k-1)p_k$ converge et exprimer sa somme en fonction de $\mathbb{E}(X)$.

(b) Justifier la convergence de la série $\sum p_k \ln(q_k)$ et démontrer que

$$\sum_{k=1}^{+\infty} p_k \ln(q_k) = \ln(p) + (\mathbb{E}(X) - 1)\ln(1-p).$$

(c) Démontrer que

$$-H(X_1) \leq \sum_{k=1}^{+\infty} p_k \ln(q_k).$$

(d) En déduire que

$$H(X) - H(X_1) \leq \sum_{k=1}^{+\infty} p_k \ln\left(\frac{q_k}{p_k}\right)$$

puis que

$$H(X) \leq H(X_1)$$

Indication : On pourra utiliser l'inégalité démontrée dans la question I.B.

Exercice 2 *(Fonction de répartition)***.**
On étudie dans cet exercice la fonction de répartition d'une var X sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Définition : La fonction de répartition de X est la fonction $F_X : \mathbb{R} \to \mathbb{R}$ définie par $F_X(x) = \mathbb{P}(X \leq x)$.

1° Résultats généraux (X suit une loi discrète quelconque). Démontrer les assertions suivantes.

a) F_X est croissante.

b) $\displaystyle\lim_{x\to-\infty} F_X(x) = 0$ et $\displaystyle\lim_{x\to+\infty} F_X(x) = 1$.

c) F_X est continue à droite en chaque réel.

d) F_X possède une limite à gauche et à droite en tout $x_0 \in \mathbb{R}$ et $\displaystyle\lim_{x\to x_0^+} F_X(x) - \lim_{x\to x_0^-} F_X(x) = \mathbb{P}(X = x_0)$.

2° Propriété fondamentale de la fonction de répartition.

a) Que dire de F_X et F_Y lorsque X et Y ont même loi ?

b) Montrer que $(F_X = F_Y)$ implique que X et Y ont même loi.

3° Une utilisation de la fonction de répartition : la loi du max de deux var indépendantes.

a) Soit X et Y deux var indépendantes sur $(\Omega, \mathcal{A}, P)$. On pose $Z = \max(X, Y)$. Démontrer que Z est une var.

b) Déterminer F_Z en fonction de F_X et F_Y.

c) Déterminer F_X puis F_Z et enfin la loi de Z lorsque $X \sim \mathcal{G}(p_1)$ et $Y \sim \mathcal{G}(p_2)$.