Samedi 15 octobre 2022 – Durée : 4h

L'usage de la calculatrice n'est pas autorisé.

Problème I. Autour des projecteurs

I.A. Un cas particulier. On considère les ensembles suivants :

$F = \{(x, y, z) \in \mathbf{R}^3 \mid 2x + y - z = 0\}$ et $G = \{(x, y, z) \in \mathbf{R}^3 \mid 2x - 2y + z = x - y - z = 0\}$

Q1. Montrer que F est un sous-espace vectoriel.

Q2. Déterminer une base de F et une base de G.

Q3. Montrer que $\mathbf{R}^3 = F \oplus G$.

Q4. Soit p la projection sur F parallèlement à G. Déterminer la matrice de p dans la base canonique de $\mathbf{R}^3$.

Q5. Justifier que la famille de vecteurs obtenue en concaténant la base de F et la base de G déterminées à la question **Q2**, est une base de $\mathbf{R}^3$. Déterminer la matrice de p dans cette base.

Dans la suite de l'exercice, E désigne un R-espace vectoriel de dimension $n \geq 2$.

I.B. Généralités. On considère un *projecteur* p de E, c'est-à-dire un endomorphisme de E tel que $p \circ p = p$.

Q6. Montrer que $E = \mathrm{Ker}(p) \oplus \mathrm{Im}(p)$. On pourra remarquer que, pour tout $x \in E$, $x = (x - p(x)) + p(x)$.

Q7. Montrer que $\mathrm{Im}(p) = \mathrm{Ker}(\mathrm{id}_E - p)$.

Q8. Justifier qu'il existe une base $\mathcal{B}$ de E telle que la matrice de p dans cette base soit diagonale, de la forme

$$\mathrm{Mat}_{\mathcal{B}}(p) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

(les coefficients non-diagonaux nuls ne sont pas représentés, par souci de lisibilité).

En déduire que $\mathrm{Tr}(p) = \mathrm{rg}(p)$.

Q9. Montrer par récurrence sur $k \in \mathbf{N}^\star$ que, si $E_1, \dots, E_k$ sont des sous-espaces vectoriels de E, alors

$$\dim(E_1 + \cdots + E_k) \leq \dim(E_1) + \cdots + \dim(E_k).$$

I.C. Condition nécessaire et suffisante pour qu'une somme de projecteurs soit un projecteur. Soit k un entier naturel supérieur ou égal à 2. On considère des projecteurs de E, notés $p_1, \ldots, p_k$, et on pose $q_k = p_1 + \cdots + p_k$.

Q10. Montrer que, si $p_i \circ p_j = 0$ pour tous $i, j \in [\![1; n]\!]$ distincts, alors q_k est un projecteur.

Dans toute la suite, on tente de montrer l'implication réciproque : on suppose que q_k est un projecteur et on souhaite montrer que $p_i \circ p_j = 0$ pour tous $i, j \in [\![1; n]\!]$ distincts.

Q11. (a) Montrer que $\mathrm{Im}(q_k) \subset \mathrm{Im}(p_1) + \cdots + \mathrm{Im}(p_k)$.

(b) À l'aide des résultats de la partie précédente, montrer que $\mathrm{rg}(q_k) = \dim\big(\mathrm{Im}(p_1) + \cdots + \mathrm{Im}(p_k)\big)$; en déduire que $\mathrm{Im}(q_k) = \mathrm{Im}(p_1) + \cdots + \mathrm{Im}(p_k)$.

(c) Montrer que $\mathrm{Im}(q_k) = \mathrm{Im}(p_1) \oplus \cdots \oplus \mathrm{Im}(p_k)$.

Q12. (a) Montrer que, pour tout $j \in [\![1; k]\!]$, $q_k \circ p_j = p_j$.

(b) En déduire que, pour tout $j \in [\![1; k]\!]$ et pour tout $x \in E$,

$$\sum_{1 \leq i \leq n,\ i \neq j} p_i(p_j(x)) = 0$$

(c) Montrer alors que, pour tout couple $(i, j) \in [\![1; k]\!]^2$ tel que $i \neq j$, on a $p_i \circ p_j = 0$.

Q13. Conclure.

(d'après BCE EDHEC S 2020)

Problème II. Entropie et « effet de surprise » en probabilités

Ce problème aborde la question de la mesure quantitative de l'information.
La première partie étudie quelques propriétés de la fonction logarithme népérien utiles pour les autres parties du problème.
La deuxième partie vise à construire les fonctions permttant de modéliser l'information contenue dans les événements de probabilité non nulle d'un espace probabilisé. **L'information contenue dans un tel événement correspond intuitivement à l'effet de surprise provoqué par la réalisation de cet événement** : un événement très surprenant contient beaucoup d'information, un événement peu surprenant contient peu d'information.
La troisième et la quatrième partie abordent respectivement la notion d'entropie d'une variable aléatoire discrète à valeurs entières, et l'entropie d'un coupe de variables aléatoires.

II.A. Questions préliminaires.

Q1. Montrer que, pour tout $x \in \mathbf{R}_+^\star$, $\ln(x) \leq x - 1$, et que $\ln(x) = x - 1$ si et seulement si $x = 1$. Interpréter graphiquement ce résultat.

Q2. 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.

II.B. Mathématisation de l'effet de surprise. Soit $(\Omega, \mathbf{P})$ un espace probabilisé fini. Si A est un événement de probabilité non nulle, on modélise la *quantité d'information contenue dans l'événement* A par

$$S(A) = f(\mathbf{P}(A)),$$

où f est une fonction $]0; 1] \to \mathbf{R}$ vérifiant les contraintes suivantes :

(i) $f(1) = 0$;

(ii) f est décroissante sur $]0;1]$;

(iii) pour tous $p, q \in]0;1]$, $f(pq) = f(p) + f(q)$;

(iv) f est continue sur $]0,1]$.

Q3. Quelle est la quantité d'information de l'événement certain ? Interpréter en terme d'effet de surprise.

Q4. Que peut-on dire de la quantité d'information contenue dans l'événement $A \cap B$ lorsque A et B sont indépendants ? Interpréter en terme d'effet de surprise.

Q5. Donner un exemple de fonction f vérifiant les quatre contraintes (i), (ii), (iii) et (iv).

Q6. Dans cette question, on se propose de déterminer l'ensemble des fonctions vérifiant ces quatre contraintes. Soit f une telle fonction.

(a) Soit $p \in]0;1]$. À l'aide d'un changement de variables, établir l'égalité

$$\frac{1}{p}\int_{p/2}^{p} f(t)\,\mathrm{d}t = \frac{1}{2}f(p) + \int_{1/2}^{1} f(u)\,\mathrm{d}u.$$

(b) En déduire que f est dérivable sur $]0;1]$.

(c) Dans cette question, on fixe $p \in]0;1]$. En dérivant par rapport à q l'égalité (iii), démontrer l'existence d'un réel a indépendant de p tel que $f'(p) = a/p$. Préciser la valeur de a.

(d) En utilisant le fait que l'égalité $f'(p) = a/p$ est valable pour tout $p\,]0;1]$, déterminer l'ensemble des fonctions vérifiant les quatre contraintes (i), (ii), (iii) et (iv).

(e) Montrer que, parmi ces fonctions, il en existe une et une seule vérifiant en plus l'égalité $f(1/e) = 1$.

Cette fonction, notée h dans la suite du problème, correspond au choix d'une unité particulière (le logon) pour mesurer la quantité d'information.

(f) Que vaut $\lim_{p\to 0} h(p)$? Interpréter ce résultat.

Q7. On réalise l'expérience aléatoire consistant à effectuer deux lancers successifs d'un dé équilibré à six faces. On considère les événements suivants :

- E : « le numéro sorti lors du premier lancer est pair » ;
- M : « le maximum des deux numéros sortis est égal à 4 » ;
- N : « la somme des deux numéros sortis est égale à 7 ».

Ordonner les quantités d'information continues dans chacun de ces trois événements. Interpréter en terme d'effet de surprise.

II.C. Entropie d'une variable aléatoire finie. Dans cette 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 $p_k = \mathbf{P}(X = k)$. 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 $p_k = 0$.

Q8. Interpréter $H(X)$ comme une espérance, puis en terme de quantité d'information.

Q9. Montrer que $H(X) \geq 0$, et que $H(X) = 0$ si et seulement si X est une variable aléatoire certaine.

Q10. Dans cette question, on considère une variable aléatoire X_0 suivant une loi uniforme sur $[\![0;n]\!]$.

(a) Calculer $H(X_0)$.

(b) En appliquant l'inégalité de la question **Q1** à un nombre réek x bien choisi, démontrer que
$$\forall\, k \in [\![0;n]\!] \quad -p_k \ln(p_k) + p_k \ln\left(\frac{1}{n+1}\right) \leq \frac{1}{n+1} - p_k.$$

(c) En déduire que, pour toute variable aléatoire $X\colon \Omega \to [\![0;n]\!]$, on a $H(X) \leq H(X_0)$, avec égalité si et seulement si X suit la même loi que X_0. Interpréter ce résultat en terme de quantité moyenne d'information.

II.D. **Entropie d'un couple de variables aléatoires finies.** Dans cette partie, m et n sont des entiers naturels non nuls. On considère deux couples de variables aléatoires (X, Y) et (X', Y'). X et X' sont à valeurs dans $[\![0;n]\!]$ et Y et Y' sont à valeurs dans $[\![0;m]\!]$. Pour tous $i \in [\![0;n]\!]$ et $j \in [\![0;m]\!]$, on note
$$p_i = \mathbf{P}(X_i), \quad q_j = \mathbf{P}(Y_j), \quad \lambda_{ij} = \mathbf{P}(X = i, Y = j) \text{ et } \lambda'_{ij} = \mathbf{P}(X' = i, Y' = j).$$
On suppose que, pour tout $(i,j) \in [\![0;n]\!] \times [\![0;m]\!]$, $\lambda_{ij} \neq 0$ et $\lambda'_{ij} \neq 0$.

On définit l'*entropie du couple* (X, Y) par
$$H(X,Y) = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij} \ln(\lambda_{ij})$$
et l'*information entre les couples* (X,Y) et (X',Y') par
$$K(X,Y,X',Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij} \ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right).$$

Q11. Rappeler les valeurs de $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij}$ et $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda'_{ij}$; en déduire que
$$K(X,Y,X',Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij}\left(\ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right) - \frac{\lambda'_{ij}}{\lambda_{ij}} + 1\right).$$

Q12. À l'aide de l'inégalité de la question **Q1**, établir que $K(X,Y,X',Y') \geq 0$, et que l'égalité a lieu si et seulement si les deux couples (X,Y) et (X',Y') ont la même loi conjointe.

Q13. On suppose que les deux variables X' et Y' sont indépendantes, que X' suit la même loi que X et que Y' suit la même loi que Y.
Démontrer que $K(X,Y,X',Y') = H(X) + H(Y) - H(X,Y)$, puis en déduire que
$$H(X,Y) \leq H(X) + H(Y).$$
Donner une condition nécessaire et suffisante pour que cette inégalité soit une égalité.

(d'après Centrale-Supélec 2017)

⋆ ⋆
⋆