Devoir surveillé 03

2TSI. Mathématiques

Mercredi 20 décembre 2017. Durée : 4 heures

Les calculatrices ou portables ou autres documents sont prohibés.
L'exercice et le problème sont indépendants et peuvent être traités dans n'importe quel ordre.

Exercice

Posé à l'oral CCP filière TSI en 2016

1. Montrer que :
$$\Phi(P, Q) = P(0)Q(0) + P'(0)Q'(0) + P''(0)Q''(0)$$
est un produit scalaire sur $\mathbb{R}_2[X]$.
2. Trouver (pour Φ) l'orthogonal de l'espace vectoriel F engendré par $X^2 + 1$ et $X^2 - X - 1$.
(Il s'agit de trouver la forme des polynômes de $\mathbb{R}_2[X]$ qui sont orthogonaux à la fois à $X^2 + 1$ et à la fois à $X^2 - X - 1$.)
3. Déterminer (pour Φ) le projeté orthogonal de X sur F.
(On pourra déterminer préalablement une base orthonormale de F par le procédé d'orthonormalisation de Gram-Schmidt.)

Problème

Extrait de l'écrit de l'épreuve de Maths I de CCS pour la filière TSI en 2017

Partie A : Questions préliminaires

1. Représenter graphiquement la fonction logarithme néperien.
2. Démontrer que, pour tout $x \in \mathbf{R}_+^\star$, $\ln(x) \leqslant x - 1$ et que $\ln x = x - 1$ si et seulement si $x = 1$.
3. Donner une interprétation graphique de ces deux résultats.
4. Montrer que la fonction g définie sur $[0, 1]$ par $g(0) = 0$ et pour tout $x \in]0, 1]$ par $g(x) = x \ln x$ est continue sur $[0, 1]$ et dérivable sur $]0, 1]$. Représenter graphiquement la fonction g.

Partie B : Mathématisation de l'effet de surprise

Soit $(\Omega, \mathcal{P})$ un espace probabilisé fini. On convient de modéliser la quantité d'information contenue dans les événements de probabilité non nulle par une fonction S définie par :

$$\forall A \in \mathcal{P}(\Omega) \text{ avec } P(A) \neq 0,\ S(A) = f(P(A)),$$

où $f : [0, 1] \to \mathbf{R}$ vérifie les contraintes suivantes :

- (i) $f(1) = 0$.
- (ii) f est décroissante sur $]0, 1]$.
- (iii) $\forall (p, q) \in]0, 1]^2$, $f(pq) = f(p) + f(q)$.
- (iv) f est continue sur $]0, 1]$.

La mesure $S(A)$ (qui est la quantification d'information contenue dans A) est considérée aussi comme la quantification de l'effet de surprise provoqué par la réalisation de cet événement A.

T.S.V.P →

1. Quelle est la quantité d'information de l'événement certain ? Interpréter en terme d'effet de surprise.
2. 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.
3. Donner un exemple de fonction f vérifiant les quatre contraintes (i), (ii), (iii) et (iv).
4. On se propose maintenant de déterminer l'ensemble des fonctions vérifiant ces quatre contraintes. Soit f une telle fonction ;

   (a) Soit $p \in]0,1]$. Établir, à l'aide d'un changement de variable, l'égalité :
   $$\frac{1}{p}\int_{\frac{p}{2}}^{p} f(t)\,dt = \frac{1}{2}f(p) + \int_{\frac{1}{2}}^{1} f(u)\,du.$$

   (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 $f'(p) = \dfrac{a}{p}$. Préciser la valeur de a.

   (d) L'égalité $f'(p) = \dfrac{a}{p}$ étant vraie quel que soit p dans $]0,1]$, déterminer l'ensemble des fonctions f 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.
   Que vaut $\lim\limits_{p\to\infty} h(p)$? Interpréter ce résultat.

   (f) On réalise l'expérience aléatoire consistant à effectuer deux lancers successifs d'un dé équilibré à six faces. On considère les évévements suivants :
   - E : « le numéro sorti lors du premier lancer est pair ».
   - « le maximum des deux numéros sortis est égal à 4 ».
   - « la somme des deux numéros sortis est égale à 7 ».

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

Partie C : Entropie d'une variable aléatoire

1. 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 aléatoire, on note $p_k = P(X=k)$. 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 $p_k = 0$.

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

   (b) Montrer que $H(X) \geqslant 0$ et que $H(X) = 0$ si et seulement si X est une variable aléatoire certaine, c'est-à-dire :
   $$\exists\, i \in [\![0,n]\!],\ p_i = 1 \text{ et } \forall j \neq i,\ p_j = 0.$$

   (c) Soit X_0 une variable aléatoire qui suit la loi uniforme sur $[\![0,n]\!]$.

      i. Calculer $H(X_0)$.

      ii. En appliquant l'inégalité de la question A-2 à un nombre réel x bien choisi, démontrer :
      $$\forall k \in [\![0,n]\!],\ -p_k \ln(p_k) + p_k \ln\left(\frac{1}{n+1}\right) \leqslant \frac{1}{n+1} - p_k.$$

      iii. En déduire que $H(X) \leqslant H(X_0)$, 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 A-2).
      Interpréter ce résultat en terme de quantité moyenne d'information.

2. Dans cette sous-partie, on s'intéresse à des variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathcal{P})$ et prenant leurs valeurs dans $\mathbf{N}^\star$.
Si X est une telle variable pour laquelle $P(X = k)$ est notée p_k, alors pour une telle variable aléatoire réelle, on a :
$$\forall k \in \mathbf{N}^\star,\ p_k \in [0,1] \text{ et } \sum_{k=1}^{+\infty} p_k = 1.$$
On dit, par ailleurs, que X est ***d'espérance finie*** si la série $\sum_{k\geqslant 1} k\,p_k$ est absolument convergente.
On dit, par ailleurs, que X est ***d'entropie finie*** si la série $\sum_{k\geqslant 1} p_k\ \ln(p_k)$ est absolument convergente et on définit alors son entropie par :
$$H(X) = -\sum_{k=0}^{+\infty} p_k \ln(p_k)$$
et on convient toujours que $p_k \ln(p_k) = 0$ vaut 0 si $p_k = 0$.
Enfin, on admet les égalités suivantes :
$$\forall |x| < 1,\ \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x} \text{ et } \sum_{k=0}^{+\infty} kx^k = \frac{x}{(1-x)^2}.$$
   (a) Pour $p \in]0,1[$ fixé, on dit que X_1 suit la loi géométrique de paramétre p si et seulement si pour tout $k \in \mathbf{N}^\star$, $P(X_1 = k) = q^{k-1}p$, où $q = 1-p$.
   Vérifier que X_1 suit bien une loi de probabilité et calculer $E(X_1)$.
   En déduire que X_1 est d'espérance finie.
   Démontrer aussi que X_1 est d'entropie finie et que $H(X_1) = -\dfrac{1-p}{p}\ln(1-p) - \ln(p)$.
   (b) Dans cette question, X est une v.a.r à espérance finie (ie $E(X) < +\infty$) et on note donc $E(X) = \sum_{k=1}^{+\infty} kP(X=k) = \sum_{k=1}^{+\infty} kp_k$. On se propose de démontrer que X est d'entropie finie.
      i. Quelle est la limite de p_k lorsque k tend vers $+\infty$?
      ii. En déduire que $\lim_{k\to+\infty} \sqrt{p_k}\ln(p_k) = 0$ puis qu'il existe un entier k_0 tel que pour tout $k \geqslant k_0$, $0 \leqslant -\sqrt{p_k}\ln(p_k) \leqslant 1$.
      iii. Soit $k \geqslant k_0$. Montrer que :
         - si $p_k \leqslant \dfrac{1}{k^3}$ alors $0 \leqslant -p_k\ln(p_k) \leqslant \dfrac{1}{k^{3/2}}$.
         - si $p_k \geqslant \dfrac{1}{k^3}$ alors $0 \leqslant -p_k\ln(p_k) \leqslant 3p_k\ln(p_k)$.
      iv. Soit $k \geqslant 1$, justifier que $\ln k \geqslant k$ puis que la série $\sum_{k\geqslant 1}\left(\dfrac{1}{k^{3/2}} + 3p_k \ln k\right)$ converge.
      v. Conclure.

Partie D : Entropie d'un couple de v.a.r et entropie conditionnelle

Dans cette partie, m et n sont des entiers non nuls, (X, Y) et (X', Y') sont deux couples de variables aléatoires discrètes, plus précisement, X et X' sont à valeurs dans $[\![0,n]\!]$, Y et Y' sont à valeurs dans $[\![0,m]\!]$. Pour $i \in [\![0,n]\!]$ et $j \in [\![0,m]\!]$, on note :
$$p_i = P(X = i),\ q_j = P(Y = j),\ \lambda_{i,j} = P(X = i, Y = j) \text{ et } \lambda'_{i,j} = P(X' = i, Y' = j).$$
On suppose que pour tout $i \in [\![0,n]\!]$ et tout $j \in [\![0,m]\!]$, $\lambda_{i,j} \neq 0$ et $\lambda'_{i,j} \neq 0$.
On définit ***l'entropie du couple*** (X, Y) par :
$$H(X,Y) = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{i,j}\ln(\lambda_{i,j}).$$

T.S.V.P $\rightarrow$

On définit ***l'information entre les couples*** (X,Y) et (X',Y') par :

$$K\left(X,Y,X',Y'\right)=-\sum_{i=0}^{n}\sum_{j=0}^{m}\lambda_{i,j}\ln\left(\frac{\lambda'_{i,j}}{\lambda_{i,j}}\right).$$

1. Propriétés de l'information entre deux couples.

   (a) Rappeler les valeurs de $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m}\lambda_{i,j}$ et de $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m}\lambda'_{i,j}$ et en déduire que :

   $$K\left(X,Y,X',Y'\right)=-\sum_{i=0}^{n}\sum_{j=0}^{m}\lambda_{i,j}\left(\ln\left(\frac{\lambda'_{i,j}}{\lambda_{i,j}}\right)-\frac{\lambda'_{i,j}}{\lambda_{i,j}}+1\right).$$

   (b) À l'aide de l'inégalité de la question A)2), établir que $K\left(X,Y,X',Y'\right)\geqslant 0$, et que l'égalité a lieu si et seulement si les deux couples (X,Y) et (X',Y') ont la même loi conjointe.

   (c) On suppose que les deux variables aléatoires X' et Y' sont indépendantes, que X' suit la même loi que X et que Y' suit la même loi que Y.
   Démontrer que $K\left(X,Y,X',Y'\right)=H(X)+H(Y)-H(X,Y)$.
   Déduire de ce qui précède que :

   $$(1)\quad H(X,Y)\leqslant H(X)+H(Y).$$

   Donner une condition nécessaire et suffisante pour que cette inégalité soit une égalité.
   Remarque : l'inégalité (1) a été obtenue en supposant, pour tout $(i,j)\in[\![0,n]\!]\times[\![0,m]\!]$, $\lambda'_{i,j}\neq 0$ et $\lambda_{i,j}\neq 0$. On admet qu'elle reste vraie même en dehors de cette condition.

2. On définit ***l'entropie conditionnelle*** de Y sachant X par :

   $$H_X(Y)=H(X,Y)-H(X).$$

   Elle mesure l'incertitude restant sur la valeur de Y lorsque la valeur de X est connue.

   (a) Montrer que $H_X(Y)\leqslant H(Y)$. Interpréter cette inégalité.

   (b) On considère $m+1$ réels a_0, a_1, ..., a_m compris entre 0 et 1.

      i. Dans cette question, on suppose $(a_0,a_1,...,a_m)\in]0,1]^{m+1}$.
      Démontrer que pour tout $j\in[\![0,m]\!]$, $\ln(a_j)\leqslant\ln(a_0+a_1+...+a_m)$.
      En déduire l'inégalité :

      $$(2)\quad \sum_{j=0}^{m}g(a_j)\leqslant g\left(\sum_{j=0}^{m}a_j\right).$$

      (La fonction g a été définie à la partie A.)

      ii. L'inégalité (2) reste-t-elle valable si $(a_0,a_1,...,a_m)\in[0,1]^{m+1}$?

      iii. Montrer que l'inégalité (2) est une égalité si et seulement s'il existe au plus un indice $j\in[\![0,m]\!]$ pour lequel $a_j\neq 0$.

   (c) Montrer que, pour tout $i\in[\![0,n]\!]$, $\displaystyle\sum_{j=0}^{m}g(\lambda_{i,j})\leqslant g(p_i)$. En déduire que $H_X(Y)\geqslant 0$.