Université Paris-Dauphine
J. Féjoz

Examen de janvier 2012

Deux heures. Sans document, ni calculatrice, ni téléphone, etc.
Chaque question numérotée vaut le même nombre de points.
Il est demandé de traiter deux des trois exercices suivants, en indiquant clairement son choix. Seuls les deux premiers exercices traités seront corrigés.

Exercice 1 Méthode de Laplace

Soit $f :]0,1[\to \mathbb{R}$ une fonction borélienne intégrable par rapport à la mesure de Lebesgue. On suppose que f possède une limite $f(1^-) \in \mathbb{R}_*$ à gauche en 1. On veut démontrer l'équivalent suivant quand n tend vers $+\infty$:

$$I_n = \int_0^1 x^n\, f(x)\, dx \sim \frac{f(1^-)}{n}.$$

1. Montrer que I_n tend vers 0 quand n tend vers $+\infty$.
2. Montrer l'équivalence voulue dans le cas où f est de classe C^1 sur $[0,1]$, en faisant une intégration par partie.
3. Pourquoi ne peut-on pas généralement appliquer le théorème de convergence dominée directement à nI_n ?
4. Démontrer l'équivalent de I_n en coupant l'intervalle $[0,1]$ en deux : un voisinage de 1 sur lequel f est bornée et où l'on pourra faire le changement de variables $y = x^n$, et son complémentaire dont la contribution est négligeable quand n tend vers l'infini.

Exercice 2 Entropie d'un ensemble fini

Soit $X = \{1, ..., r\}$ ($r > 0$) muni d'une probabilité P. On notera $p_x = P(\{x\})$ si $x \in X$. On appelle *entropie* de X le réel positif

$$h = h(p_1, ..., p_r) = -\sum_{x\in X} p_x \ln p_x,$$

avec la convention $0 \ln 0 = 0$ dans les cas où $p_x = 0$.

1. Montrer que h atteint son maximum quand $p_x = 1/r$ pour tout $x = 1, ..., r$, d'abord en supposant que $p_x > 0$ pour tout x (on pourra maximiser h sous la contrainte $p_1 + \cdots + p_r = 1$), puis dans le cas général.

Soit $\Omega = X^n$ l'espace des suites de n tirages indépendants dans X, muni de la probabilité produit. Soit $\nu_x : \Omega \to \{0, ..., n\}$ la variable aléatoire définie par $\nu_x(\omega) =$ le nombre de composantes de $\omega = (\omega_1, ..., \omega_n)$ égales à x.

2. Calculer la loi de ν_x en fonction des p_x, $x \in X$.
3. Un nombre $\delta > 0$ étant fixé, posons

$$C_n = \{\omega,\ |\nu_x(\omega)/n - p_x| \leq \delta\ (\forall x \in X)\}.$$

   Quelle est la limite de $P(C_n)$ quand n tend vers l'infin ? On pourra utiliser la loi faible des grands nombres.
4. En déduire que, si $\alpha > 0$ est donné et si δ est donné assez petit en fonction de α, pour tout entier n et pour tout $\omega \in C_n$,

$$e^{-n(h+\alpha)} \leq P(\{\omega\}) \leq e^{-n(h-\alpha)}.$$

5. En déduire enfin le théorème de McMillan : il existe $\beta > 0$ tel que, si n est assez grand, le cardinal de C_n satisfait

$$e^{n(h-\beta)} \leq |C_n| \leq e^{n(h+\beta)}.$$

   Au regard des deux dernières questions, interpréter la place de C_n dans Ω.

Exercice 3 Un exemple de vecteur gaussien

Soient X_1, X_2 et X_3 trois variables aléatoires réelles inhdépendantes de loi $N_1(0,1)$.

1. Montrer que $X_1 + X_2$ et $X_1 - X_2$ sont indépendantes.

Notons Y le vecteur aléatoire $Y = AX$, où $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ et $X = (X_1, X_2, X_3)$.

2. Calculer la loi de Y (on pourra déterminer d'abord sa fonction caractéristique) puis, si elle existe, sa densité par rapport à la mesure de Lebesgue.
3. Constuire un vecteur aléatoire Z à valeurs dans $\mathbb{R}^2$ et à composantes indépendantes, de la forme $Z = PY$ avec $P \in Gl_2(\mathbb{R})$; on pourra commencer par diagonaliser D.

Notons $\tilde{X} = (X_1, X_2)$. On rappelle que l'espérance conditionnelle $E(X_3|\tilde{X})$ est la projection orthogonale de X_3 sur l'espace $L^2(\sigma(\tilde{X}))$, où $\sigma(\tilde{X})$ est la plus petite tribu rendant le vecteur aléatoire $\tilde{X}$ borélien.

4. Montrer qu'une variable aléatoire réelle W est $\sigma(\tilde{X})$-mesurable si et seulement si il existe une fonction $h : \mathbb{R}^2 \to \mathbb{R}$ borélienne telle que $W = h \circ \tilde{X}$.
5. Montrer que dans le cas où $W = E(X_3|X_1, X_2)$, W est la projection orthogonale V de X_3 sur le plan vectoriel réel engendré par X_1 et X_2 (et donc h est une forme linéaire) ; on pourra commencer par justifier que $X_3 - V$ est indépendante de $\tilde{X}$ et que donc $E(X_3 - V|\tilde{X}) = E(X_3 - V) = 0$.

Solution de l'exercice 1

1. Pour tout $x \in [0,1[$, la suite numérique $(x^n f(x))_{n\in\mathbb{N}}$ tend vers 0. Donc presque partout sur $[0,1]$ la suite de fonctions $(x^n f(x))_{n\in\mathbb{N}}$ converge simplement vers 0. Par ailleurs, pour tout $n \in \mathbb{N}$ la fonction $x \mapsto x^n f(x)$ est dominée sur $[0,1]$, en valeur absolue, par la fonction intégrable f. Donc d'après le théorème de convergence dominée l'intégrale I_n tend vers 0 quand n tend vers $+\infty$.
2. Une intégration par parties montre que
$$nI_n = \frac{n}{n+1}\left(f(1) - \int_0^1 x^{n+1} f'(x)\,dx\right).$$
Le premier terme du membre de droite de l'égalité tend vers $f(1)$. Le second tend, lui, vers 0, comme une simple application du théorème de convergence dominée le montre (remplacer f par f' et n par $n+1$ dans la question précédente). Finalement, $nI_n \to f(1)$, et l'équivalent en découle.
3. Quand n tend vers $+\infty$, la suite de fonctions $(f_n(x))_{n\in\mathbb{N}} = (nx^n f(x))_{n\in\mathbb{N}}$ tend simplement vers 0 presque partout. Pour appliquer le théorème de convergence dominée, il faudrait de plus montrer que la suite $(f_n)_n$ est majorée par une fonction intégrable, presque partout sur $[0,1]$.
Mais une telle fonction intégrable, en général, n'existe pas. En effet, si elle existait, on pourrait appliquer le théorème de convergence dominée et conclure que nI_n tend vers 0 quand n tend vers $+\infty$. Or la question (2) prouve que ce n'est pas le cas, déjà quand f est classe C^1.
4. Comme f a une limite finie quand x tend vers 1, il existe deux réels $\alpha \in [0,1[$ et $M > 0$ tels que $|f(x)| \leq M$ sur $[\alpha, 1]$.
Notons $f_n(x) = nx^n f(x)$. Quand n tend vers $+\infty$, la suite (f_n) tend simplement vers 0 sur $[0,\alpha]$. En outre, comme nx^n tend vers 0, on a $|f_n(x)| \leq |f(x)|$ sur $[0,\alpha]$ à partir d'un certain rang. Donc, d'après le théorème de convergence dominée,
$$\int_0^\alpha nx^n f(x)\,dx \to 0.$$
Par ailleurs on a
$$\int_\alpha^1 nx^n f(x)\,dx = \int_{\alpha^n}^1 y^{1/n} f(y^{1/n})\,dy.$$
Or la suite des fonctions $y^{1/n} f(y^{1/n}) \mathbb{1}_{[\alpha^n,1]}$ tend simplement vers la fonction constante $f(1^-)$ et elle est majorée en valeur absolue par la fonction constante M, qui est intégrable sur $[0,1]$. Donc, d'après le théorème de convergence dominée,
$$\int_\alpha^1 nx^n f(x)\,dx \to \int_0^1 f(1^-)\,dy = f(1^-).$$
Finalement, $nI_n \to f(1^-)$, et, si $f(1^-) \neq 0$,
$$\int_0^1 x^n f(x)\,dx \sim \frac{f(1^-)}{n}.$$

Solution de l'exercice 2

1. Soit $(p_1, ..., p_r)$ un maximum local de $h(p_1, ..., p_r)$ sous la contrainte $p_1 + \cdots + p_r = 1$.
   Si toutes les probabilités p_x sont > 0, d'après le principe d'extrémalisation de Lagrange, il existe un réel λ tel que
   $$dh = \lambda(dp_1 + \cdots + dp_r),$$
   soit
   $$\frac{\partial h}{\partial p_x} = -\ln p_x - 1 = \lambda \quad (\forall x \in X),$$
   soit
   $$p_x = e^{-\lambda-1} \quad (\forall x \in X) \ ;$$
   donc les probabilités p_x sont indépendantes de x, donc forcément égales chacune à $1/r$. Dans ce cas, l'entropie vaut
   $$h = \sum_x \frac{1}{r} \ln r = \ln r.$$
   Si seulement certaines probabilités p_x sont non nulles, disons s parmi r (et comme $p_1 + \cdots + p_r = 1$, on a $s \geq 1$), le même raisonnement que précédemment montre que $h = \ln s \leq \ln r$.
   Finalement, le maximum de h est bien atteint quand $p_x \equiv \frac{1}{r}$, et h vaut alors $\ln r$.
2. La loi de ν_x est binomiale : $P(\nu_x = k) = C_n^k p_x^k (1 - p_x)^{n-k}$, $k = 0, ..., n$.
3. Fixons d'abord $x \in X$. On a
   $$\nu_x(\omega) = \sum_{1 \leq j \leq n} \mathbb{1}_{\{x\}}(\omega_j), \quad \omega = (\omega_1, ..., \omega_n) \in \Omega.$$
   Donc, d'après la loi faible des grands nombre, quand n tend vers l'infini, ν_x/n tend en probabilité vers $E(\mathbb{1}_{\{x\}}) = p_x$. C'est dire que, δ étant fixé, la probabilité de
   $$C_{n,x} = \{\omega, \ |\nu_x(\omega)/n - p_x| \leq \delta\}$$
   tend vers 1. Donc, si l'on considère maintenant l'ensemble
   $$C_n = \{\omega, \ |\nu_x(\omega)/n - p_x| \leq \delta \ (\forall x \in X)\} = \cap_{x \in X} C_{n,x},$$
   on obtient
   $$P(\complement C_n) = P\left(\cup_{x \in X} \complement C_{n,x}\right) \leq \sum_{x \in X} P(\complement C_{n,x}) \to 0.$$
   Donc $P(C_n)$ tend vers 1.

4. Soit N tel que $P(C_n) \geq 1-\alpha$ pour $n \geq N$. Quel que soit $\omega = (\omega_1, ..., \omega_n) \in \Omega$,

$$\begin{aligned} P(\{\omega\}) &= \prod_{1 \leq s \leq n} p_{\omega_s} = \prod_{1 \leq x \leq r} p_x^{\nu_x(\omega)} \\ &= \exp\left(\sum_{1 \leq x \leq r} \nu_x(\omega) \ln p_x \right). \end{aligned}$$

Comme on veut comparer $P(\{\omega\})$ à $e^{-nh} = e^{n \sum_x p_x \ln x}$, faisons artificiellement apparaître h dans l'expression précédente, en écrivant :

$$P(\{\omega\}) = \exp\left(n \left(h + \sum_{1 \leq x \leq r} \left(\frac{\nu_x(\omega)}{n} - p_x \right) \ln p_x \right) \right).$$

Si $\omega \in C_n$,

$$\left| \sum_{1 \leq x \leq r} \left(\frac{\nu_x(\omega)}{n} - p_x \right) \ln p_x \right| \leq \delta \sum_{1 \leq x \leq n} |\ln p_x|.$$

En choisissant δ de sorte que ce dernier majorant soit $\leq \alpha$, on voit que

$$e^{-n(h+\alpha)} \leq P(\{\omega\}) \leq e^{-n(h-\alpha)}.$$

5. Sommons l'estimation précédente sur tous les $\omega \in C_n$:

$$|C_n| e^{-n(h+\alpha)} \leq \sum_{\omega \in C_n} P(\{\omega\}) = P(C_n) \leq |C_n| e^{-n(h-\alpha)},$$

donc

$$P(C_n) e^{n(h-\alpha)} \leq |C_n| \leq P(C_n) e^{n(h+\alpha)} \leq e^{n(h+\alpha)}.$$

Pour minorer $|C_n|$ indépendemment de $P(C_n)$, on se rappelle que $P(C_n)$ tend vers 1 quand n tend vers l'infini. Donc, si $\beta > \alpha$ et si n est assez grand,

$$P(C_n) \geq e^{-n(\beta-\alpha)}.$$

Alors on a bien

$$e^{n(h-\beta)} \leq |C_n| \leq e^{n(h+\alpha)} \leq e^{n(h+\beta)}.$$

Donc le cardinal de C_n est négligeable devant celui de Ω ($|\Omega| = r^n$). Pourtant, C_n concentre la quasi-totalité des issues possibles, et de plus en plus au fur et à mesure que n augmente, puisque $P(C_n) \to 1$.

Solution de l'exercice 3

1. On a

$$\mathrm{Cov}\,(U, V) = E((X_1 + X_2)(X_1 - X_2)) = E(X_1^2 - X_2^2) = 0,$$

donc les variables gaussiennes $U = X_1 + X_2$ et $V = X_1 - X_2$ sont indépendantes.

2. La fonction caractéristique de X est
$$\varphi_X(s) = e^{-\frac{1}{2}\|s\|^2} \quad (s \in \mathbb{R}^3),$$
donc celle de $Y = AX$ vaut
$$\varphi_Y(t) = \varphi_X(A' \cdot t) = e^{-\frac{1}{2}t' \cdot (AA') \cdot t} \quad (\forall t \in \mathbb{R}^2),$$
donc, par injectivité de la transformation de Fourier, la loi de Y est $N_2(0, D)$, avec $D = AA' = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Comme la matrice de dispersion D de Y n'est pas diagonale, Y n'est pas un vecteur aléatoire indépendant. Comme $\det D = 3$, D est inversible, la densité de Y est
$$\frac{1}{2\pi\sqrt{3}} \exp\left(-\frac{1}{2}y' \cdot D^{-1} \cdot y\right), \quad D^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$
soit
$$\frac{1}{2\pi\sqrt{3}} \exp\left(\frac{1}{3}\left(y_1^2 - y_1 y_2 + y_2^2\right)\right).$$
3. Posons $u = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $v = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. On voit que $Du = u$ et $Dv = 3v$. Donc, si $P = (u, v)$ est la matrice dont les colonnes sont les vecteurs u et v, P est orthogonale et l'on a
$$E := P^{-1}DP = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$
Considérons le vecteur aléatoire $Z = P^{-1}Y$, $P^{-1} = P' = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Par le même raisonnement qu'à la question précédente, la matrice de dispersion de Z est $P^{-1}DP^{-1\prime} = P'DP = E$, donc Z répond à la question posée.
4. Voir le théorème 5.7 du cours polycopié ou le dernier exercice du partiel.
5. Notons $V = \lambda_1 X_1 + \lambda_2 X_2$ la projection orthogonale de X_3 sur le plan vectoriel engendré par X_1 et X_2. Remarquons que
$$\operatorname{cov}(X_3 - V, X_i) = E[(X - V)X_i] = 0, \quad i = 1, 2,$$
(par définition de la projection orthogonale), donc $X - V$ est indépendant de $\tilde{X}$, donc
$$E(X_3 - V|\tilde{X}) = E(X_3 - V).$$
(Cette dernière égalité vient du fait que, si T est $\sigma(\tilde{X})$-mesurable, T est indépendante de $X_3 - V$, donc $E((X_3 - V)T) = E(X_3 - V)E(T) = E(E(X_3 - V)T)$.) Donc
$$E(X_3 - V|\tilde{X}) = E(X_3) - E(V) = 0 - 0 = 0.$$
L'espérance conditionnelle cherchée est
$$W = E(X_3|\tilde{X}) = E(X_3 - V|\tilde{X}) + E(V|\tilde{X}) = E(V|\tilde{X}).$$
Comme de plus V est $\sigma(\tilde{X})$-mesurable, $E(V|\tilde{X}) = V$. Donc $W = V$.