



PRÉPASIMPLE

prepasimple.fr

EM Lyon

Mathématiques — Option économique

2013

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Corrigé

Exercice 1

Partie I - Calcul d'une intégrale dépendant d'un paramètre

1. Sur $]0; 1]$, la fonction $t \mapsto -t \ln t$ est continue par produit de fonctions continues. Par croissance comparée, $t \ln t \rightarrow 0$ lorsque $t \rightarrow 0^+$: on peut écrire $-t \ln t = \ln u/u$ avec $u = 1/t \rightarrow +\infty$.

Cette limite vaut $g(0) = 0$. Ainsi g est continue sur tout le segment $[0; 1]$.

2. Pour $0 < x < 1$, les fonctions $t \mapsto \ln t$ et $t \mapsto t^2/2$ sont de classe C^1 sur $[x; 1]$.
L'intégration par parties donne :

$$\int_x^1 -t \ln t \, dt = \left[-\frac{t^2}{2} \ln t \right]_x^1 + \frac{1}{2} \int_x^1 t \, dt.$$

Par conséquent :

$$\int_x^1 g(t) \, dt = \frac{x^2}{2} \ln x + \frac{1 - x^2}{4}.$$

3. La continuité de g sur $[0; 1]$ assure déjà l'existence de son intégrale ordinaire. Par I.2 et la croissance comparée $x^2 \ln x \rightarrow 0$ lorsque $x \rightarrow 0^+$:

$$\int_0^1 g(t) \, dt = \lim_{x \rightarrow 0^+} \left(\frac{x^2}{2} \ln x + \frac{1 - x^2}{4} \right) = \frac{1}{4}.$$

Il n'y a donc aucune singularité non intégrable à traiter à l'origine : g y a été prolongée continûment.

Partie II - Exemple de densité

1. Sur $]0; 1[$, f est une somme de fonctions continues. Sur $] - \infty; 0[$ et $]1; +\infty[$, elle est nulle, donc continue. En 0, les deux limites valent 0, puisque $-t \ln t \rightarrow 0$ et $t^{1/3} \rightarrow 0$ à droite; elles sont égales à $f(0)$.

Ainsi f est continue sur $] - \infty; 1[$ et sur $]1; +\infty[$. En revanche :

$$\lim_{t \rightarrow 1^-} f(t) = 1, \quad f(1) = 0, \quad \lim_{t \rightarrow 1^+} f(t) = 0.$$

Elle n'est pas continue en 1; ses limites latérales y sont finies.

2. Les intégrales hors de $[0; 1]$ sont nulles. Sur $[0; 1]$, g et $t \mapsto t^{1/3}$ sont continues; la valeur assignée à f en 1 ne change pas l'intégrale. Les intégrales existent et :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^1 g(t) dt + \int_0^1 t^{1/3} dt = \frac{1}{4} + \left[\frac{3}{4} t^{4/3} \right]_0^1 = 1.$$

3. Si $0 < t < 1$, alors $-t \ln t \geq 0$ et $t^{1/3} > 0$; ailleurs $f = 0$. Ainsi f est positive sur $\mathbb{R}$. Elle est continue sauf au seul point 1, possède des limites latérales finies en tout point et son intégrale sur $\mathbb{R}$ vaut 1. Ce sont les conditions suffisantes pour qu'elle soit une densité de probabilité.
4. a. Sur $]0; 1[$, le logarithme et la fonction puissance considérée sont de classe C^2 . Leurs dérivées donnent :

$$f'(t) = -\ln t - 1 + \frac{1}{3t^{2/3}}, \quad f''(t) = -\frac{1}{t} - \frac{2}{9t^{5/3}}.$$

En particulier, $f''(t) < 0$ pour tout $t \in]0; 1[$.

b. La fonction f' est continue et strictement décroissante sur $]0; 1[$. Elle vérifie :

$$f'(1/e) = \frac{e^{2/3}}{3} > 0, \quad \lim_{t \rightarrow 1^-} f'(t) = -\frac{2}{3} < 0.$$

On peut prolonger la formule de f' continûment jusqu'en 1 en lui attribuant la valeur $-2/3$; cette valeur de la formule ne désigne pas une dérivée de f en son point de discontinuité. Le théorème des valeurs intermédiaires, puis la stricte décroissance, donnent une unique racine $\alpha \in]1/e; 1[$. Sur $]0; 1/e]$, f' reste positive ; il n'existe donc aucune autre racine dans $]0; 1[$.

c. Question adaptée au programme actuel. Code adapté en Python.

On conserve un intervalle $[a; b]$ contenant α , avec $h(a) \geq 0$ et $h(b) \leq 0$, où $h(t) = -\ln t - 1 + 1/(3t^{2/3})$. Sa longueur est divisée par deux à chaque tour. Le milieu final approche α avec une erreur au plus égale à la demi-longueur de l'intervalle.

Python

```
import numpy as np

def h(t):
    return -np.log(t) - 1 + 1 / (3 * t ** (2 / 3))

a = np.exp(-1)
b = 1.0
while b - a > 2 * 10 ** (-3):
    m = (a + b) / 2
    if h(m) > 0:
        a = m
    else:
        b = m
alpha_approchee = (a + b) / 2
print(alpha_approchee)
```

Le programme termine car la longueur initiale $1 - e^{-1}$ est divisée par 2 à chaque

itération. Il affiche environ 0,5907 ; l'erreur est inférieure à 10^{-3} . Le code complet a été exécuté ; l'encadrement et la précision résultent de la dichotomie, et ne dépendent pas d'une simple observation numérique.

Partie III - Calcul d'une fonction de répartition

1. Pour $0 < x < 1$, la linéarité de l'intégrale et I.2 donnent :

$$\int_x^1 f(t) dt = \frac{x^2}{2} \ln x + \frac{1-x^2}{4} + \frac{3}{4}(1-x^{4/3}).$$

Ainsi :

$$\int_x^1 f(t) dt = 1 + \frac{x^2}{2} \ln x - \frac{x^2}{4} - \frac{3}{4}x^{4/3}.$$

2. Par définition de la fonction de répartition d'une variable à densité,

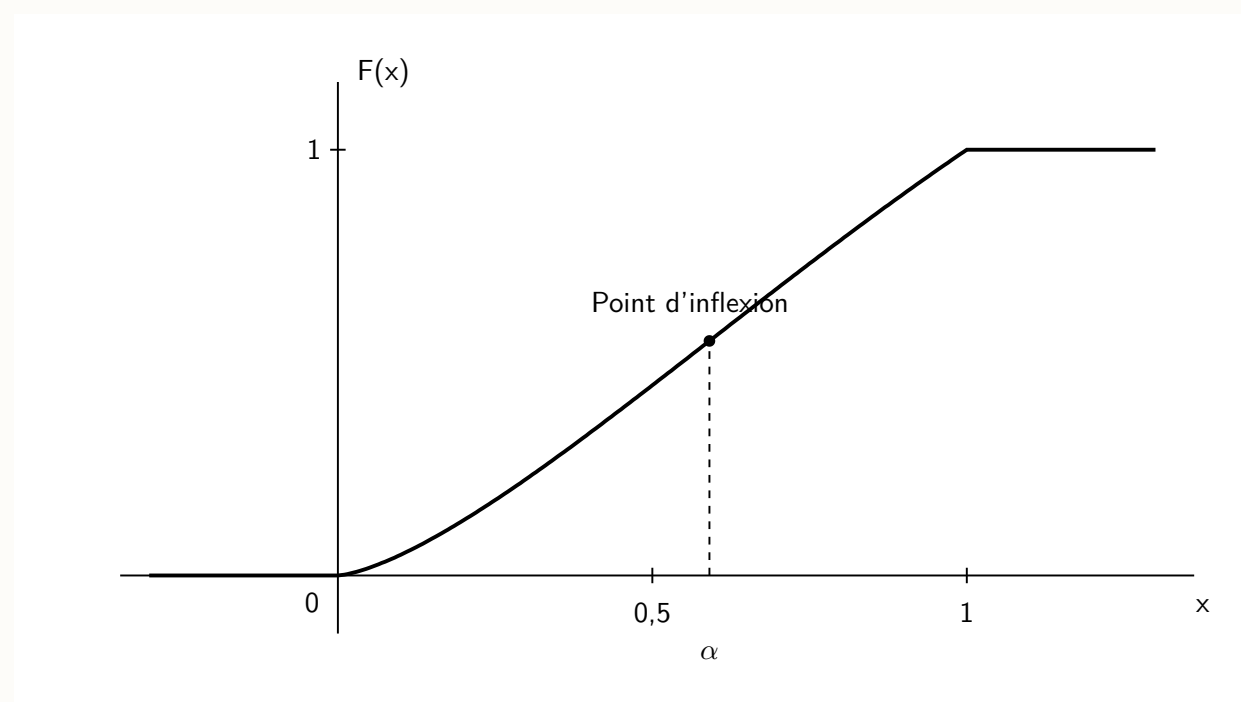
$F(x) = \int_{-\infty}^x f(t) dt$. Si $x \leq 0$, cette intégrale est nulle. Si $x \geq 1$, elle vaut 1. Enfin, pour $0 < x < 1$, on utilise la normalisation et III.1 :

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \frac{x^2}{4} - \frac{x^2}{2} \ln x + \frac{3}{4}x^{4/3} & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

La limite à droite en 0 est 0, et la limite à gauche en 1 est 1. Cette expression est donc continue sur $\mathbb{R}$, comme attendu pour une variable à densité.

3. Sur $]0; 1[$, $F' = f > 0$: la fonction est strictement croissante. Comme $F'' = f'$, elle est convexe sur $]0; \alpha[$ et concave sur $]\alpha; 1[$. Elle possède un point d'inflexion d'abscisse α .

La courbe est horizontale d'ordonnée 0 pour $x \leq 0$, puis rejoint $(1; 1)$ et devient horizontale d'ordonnée 1 pour $x \geq 1$. En 0, la tangente est horizontale : en divisant l'expression de $F(x)$ par $x > 0$, on obtient une limite nulle. En 1, le taux d'accroissement à gauche tend vers 1, tandis que celui de droite est nul : F n'est pas dérivable en 1.

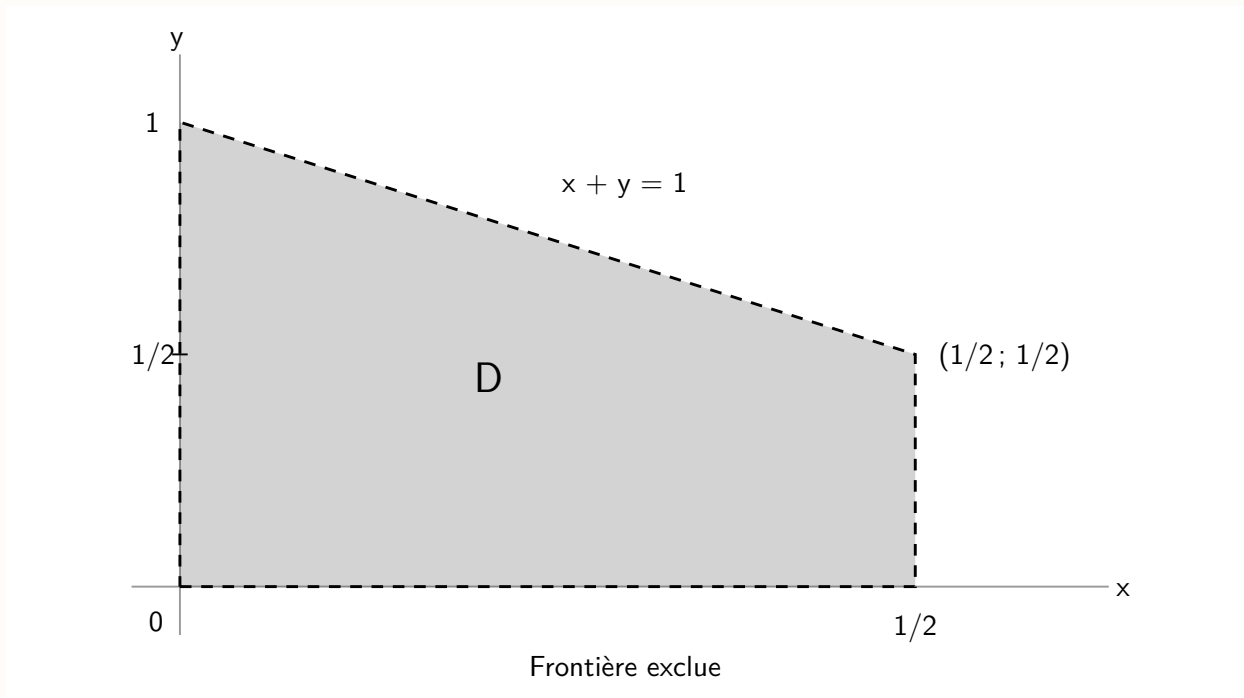


Partie IV - Étude d'extremum local pour une fonction de deux variables réelles

1. Les conditions définissant D s'écrivent :

$$0 < x < \frac{1}{2}, \quad 0 < y < 1 - x.$$

Il s'agit de l'intérieur du quadrilatère de sommets $(0; 0)$, $(1/2; 0)$, $(1/2; 1/2)$ et $(0; 1)$; aucun côté ni sommet n'appartient à D . La zone grisée ci-dessous représente cet ouvert, les pointillés sa frontière exclue.



2. Pour tout $(x, y) \in D$, on a $x + y \in]0; 1[$ et $2x \in]0; 1[$, ce qui permet de dériver les deux compositions. La règle de la chaîne donne :

$$\frac{\partial G}{\partial x}(x, y) = f'(x + y) - f'(2x), \quad \frac{\partial G}{\partial y}(x, y) = f'(x + y).$$

3. Le gradient s'annule si et seulement si $f'(x + y) = 0$ et $f'(x + y) - f'(2x) = 0$. En substituant la première égalité dans la seconde, cela équivaut à :

$$f'(2x) = 0 \quad \text{et} \quad f'(x + y) = 0.$$

La réciproque résulte immédiatement des formules de IV.2.

4. Les deux arguments appartiennent à $]0; 1[$ et la seule racine de f' sur cet intervalle est α . Les équations de IV.3 équivalent donc à $2x = \alpha$ et $x + y = \alpha$, soit :

$$(x, y) = \left(\frac{\alpha}{2}, \frac{\alpha}{2} \right).$$

Ce point appartient à D : ses coordonnées sont strictement positives et $x + y = 2x = \alpha < 1$. Il est donc bien l'unique point critique de G .

5. La fonction G est de classe C^2 sur l'ouvert D , puisque les arguments $x + y$ et $2x$ restent dans $]0; 1[$. Sa hessienne est :

$$\nabla^2 G(x, y) = \begin{pmatrix} f''(x + y) - 2f''(2x) & f''(x + y) \\ f''(x + y) & f''(x + y) \end{pmatrix}.$$

Au point critique, posons $c = f''(\alpha) < 0$. Il vient :

$$H = \nabla^2 G(\alpha/2, \alpha/2) = c \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

On détermine les valeurs propres sans déterminant. Pour $B = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$, les produits explicites donnent :

$$B \begin{pmatrix} 1 \\ 1 + \sqrt{2} \end{pmatrix} = \sqrt{2} \begin{pmatrix} 1 \\ 1 + \sqrt{2} \end{pmatrix}, \quad B \begin{pmatrix} 1 \\ 1 - \sqrt{2} \end{pmatrix} = -\sqrt{2} \begin{pmatrix} 1 \\ 1 - \sqrt{2} \end{pmatrix}.$$

Ces deux vecteurs sont indépendants, car associés à deux valeurs propres distinctes. Les valeurs propres de H sont donc $c\sqrt{2} < 0$ et $-c\sqrt{2} > 0$. Le critère par la hessienne montre que l'unique point critique est un point selle. Il n'est pas un extremum local.

Tout extremum local d'une fonction de classe C^1 sur un ouvert est un point critique ; il n'en existe aucun autre. Ainsi G n'admet aucun extremum local sur D .

Exercice 2

1. La matrice A est réelle et symétrique : sa transposée est elle-même. Le résultat du programme sur les matrices symétriques réelles assure qu'elle est diagonalisable dans $\mathcal{M}_4(\mathbb{R})$.
2. Notons (e_1, e_2, e_3, e_4) la base canonique. Les colonnes de A donnent $Ae_1 = 2e_4$, $Ae_4 = 2e_1$, $Ae_2 = e_3$ et $Ae_3 = e_2$. Par addition et soustraction :

$$A(e_1 - e_4) = -2(e_1 - e_4), \quad A(e_2 - e_3) = -(e_2 - e_3),$$

$$A(e_2 + e_3) = e_2 + e_3, \quad A(e_1 + e_4) = 2(e_1 + e_4).$$

Ces quatre vecteurs non nuls sont associés à quatre valeurs propres distinctes ; ils sont libres et forment donc une base de $\mathbb{R}^4$. Dans cette base, A est diagonale avec les quatre coefficients $-2, -1, 1, 2$. Il ne peut donc y avoir d'autre valeur propre et chaque sous-espace propre est de dimension 1. Précisément :

$$\text{Sp}(A) = \{-2, -1, 1, 2\},$$

$$E_{-2} = \text{Vect}(e_1 - e_4), \quad E_{-1} = \text{Vect}(e_2 - e_3), \quad E_1 = \text{Vect}(e_2 + e_3), \quad E_2 = \text{Vect}(e_1 + e_4).$$

3. On range les vecteurs propres précédents dans l'ordre croissant des valeurs propres. Cela donne :

$$D = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

Les quatre coefficients diagonaux de P valent 1. Ses colonnes forment une base, donc elle est inversible. L'égalité $AP = PD$ donne $A = PDP^{-1}$.

Pour calculer l'inverse, si $X = PY$, les coordonnées vérifient $x_1 = y_1 + y_4$, $x_2 = y_2 + y_3$, $x_3 = -y_2 + y_3$ et $x_4 = -y_1 + y_4$. Ainsi :

$$y_1 = \frac{x_1 - x_4}{2}, \quad y_2 = \frac{x_2 - x_3}{2}, \quad y_3 = \frac{x_2 + x_3}{2}, \quad y_4 = \frac{x_1 + x_4}{2}.$$

On en déduit :

$$P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

4. La matrice nulle commute avec A . Si $M_1, M_2 \in C_A$ et $u, v \in \mathbb{R}$, alors, par distributivité des produits matriciels :

$$A(uM_1 + vM_2) = uAM_1 + vAM_2 = uM_1A + vM_2A = (uM_1 + vM_2)A.$$

L'ensemble C_A contient donc la matrice nulle et est stable par combinaison linéaire :

c'est un sous-espace vectoriel de $\mathcal{M}_4(\mathbb{R})$.

5. Puisque $A = PDP^{-1}$, on a $P^{-1}AP = D$. Multiplier l'égalité $AM = MA$ à gauche par P^{-1} et à droite par P conduit à :

$$P^{-1}AMP = P^{-1}MAP \iff D(P^{-1}MP) = (P^{-1}MP)D.$$

Toutes ces opérations sont réversibles puisque P est inversible. Par définition de N , on a donc $M \in C_A$ si et seulement si $DN = ND$, c'est-à-dire $N \in C_D$.

6. Écrivons $D = \text{diag}(d_1, d_2, d_3, d_4)$ et $N = (n_{ij})$. Les coefficients de DN et ND sont respectivement $d_i n_{ij}$ et $n_{ij} d_j$. Ainsi :

$$DN = ND \iff \forall i, j, (d_i - d_j)n_{ij} = 0.$$

Les quatre d_i sont distincts. Pour $i \neq j$, on a donc $n_{ij} = 0$, tandis que les coefficients diagonaux sont libres. Réciproquement, deux matrices diagonales commutent. Par conséquent :

$$C_D = \{\text{diag}(u, v, w, z); (u, v, w, z) \in \mathbb{R}^4\}.$$

7. Méthode 1 — Retour à la base canonique

Les questions 5 et 6 donnent $C_A = \{PNP^{-1}; N \text{ diagonale}\}$. En utilisant les matrices explicites de la question 3 :

$$P \operatorname{diag}(u, v, w, z) P^{-1} = \begin{pmatrix} \frac{u+z}{2} & 0 & 0 & \frac{z-u}{2} \\ 0 & \frac{v+w}{2} & \frac{w-v}{2} & 0 \\ 0 & \frac{w-v}{2} & \frac{v+w}{2} & 0 \\ \frac{z-u}{2} & 0 & 0 & \frac{u+z}{2} \end{pmatrix}.$$

Posons $a = (u+z)/2$, $b = (z-u)/2$, $c = (v+w)/2$ et $d = (w-v)/2$. Tous les quadruplets (a, b, c, d) sont possibles : l'inverse de ce changement de paramètres est $u = a - b$, $z = a + b$, $v = c - d$, $w = c + d$. D'où exactement :

$$C_A = \left\{ \begin{pmatrix} a & 0 & 0 & b \\ 0 & c & d & 0 \\ 0 & d & c & 0 \\ b & 0 & 0 & a \end{pmatrix} ; (a, b, c, d) \in \mathbb{R}^4 \right\}.$$

Méthode 2 — Action sur les vecteurs propres

Cette méthode explique directement pourquoi la diagonalisation simplifie la commutation. Si $AM = MA$ et $Av = \lambda v$, alors $A(Mv) = M(Av) = \lambda Mv$. Donc M laisse chaque droite propre de A stable. Réciproquement, si M laisse chacune des quatre droites propres stable, alors AM et MA coïncident sur les quatre vecteurs d'une base, donc sont égales.

Il existe donc u, v, w, z tels que $M(e_1 - e_4) = u(e_1 - e_4)$, $M(e_2 - e_3) = v(e_2 - e_3)$, $M(e_2 + e_3) = w(e_2 + e_3)$, $M(e_1 + e_4) = z(e_1 + e_4)$. En additionnant et en soustrayant ces égalités, on obtient :

$$Me_1 = ae_1 + be_4, \quad Me_4 = be_1 + ae_4, \quad Me_2 = ce_2 + de_3, \quad Me_3 = de_2 + ce_3,$$

avec les mêmes paramètres que dans la première méthode, arbitraires. Les colonnes de M ont donc la forme annoncée, qui suffit également à assurer la commutation.

8. Notons E_{ij} la matrice dont l'unique coefficient non nul est le coefficient (i, j) , égal à 1. Les matrices :

$$B_1 = E_{11} + E_{44}, \quad B_2 = E_{14} + E_{41}, \quad B_3 = E_{22} + E_{33}, \quad B_4 = E_{23} + E_{32}$$

engendrent C_A d'après sa description à la question 7. Si $aB_1 + bB_2 + cB_3 + dB_4 = 0$, la lecture des coefficients $(1, 1)$, $(1, 4)$, $(2, 2)$ et $(2, 3)$ donne respectivement $a = b = c = d = 0$. Cette famille est donc libre, et constitue une base de C_A . Ainsi $\dim C_A = 4$.

Exercice 3

Les tirages avec remise sont indépendants et uniformes sur les n numéros. Toutes les variables considérées dans cet exercice prennent un nombre fini de valeurs ; leurs espérances, variances et covariances existent donc.

Partie I

1. Pour un numéro i fixé, chaque tirage donne la boule i avec probabilité $1/n$, indépendamment des autres tirages. Le nombre d'obtentions de i en k tirages suit donc :

$$X_i \sim \mathcal{B}(k, 1/n).$$

Pour tout $r \in \llbracket 0; k \rrbracket$, $P(X_i = r) = \binom{k}{r} (1/n)^r (1 - 1/n)^{k-r}$. Les moments d'une loi binomiale donnent :

$$E(X_i) = \frac{k}{n}, \quad V(X_i) = \frac{k}{n} \left(1 - \frac{1}{n}\right) = \frac{k(n-1)}{n^2}.$$

2. Les événements $[X_1 = k]$ et $[X_2 = k]$ ont chacun la probabilité $(1/n)^k > 0$. Ils sont incompatibles car $k \geq 1$: les k premiers tirages ne peuvent tous porter simultanément les deux numéros distincts 1 et 2. Ainsi :

$$P(X_1 = k, X_2 = k) = 0 \neq P(X_1 = k)P(X_2 = k).$$

Les variables X_1 et X_2 ne sont pas indépendantes; a fortiori la famille $X_1, \dots, X_n$ n'est pas indépendante.

3. On fixe deux numéros distincts i et j .

- a. La variable $X_i + X_j$ compte les tirages dont le résultat appartient à la paire $\{i, j\}$. Chaque tirage réalise cet événement avec probabilité $2/n$, et les tirages sont indépendants. Par conséquent :

$$X_i + X_j \sim \mathcal{B}(k, 2/n), \quad V(X_i + X_j) = \frac{2k}{n} \left(1 - \frac{2}{n}\right) = \frac{2k(n-2)}{n^2}.$$

Si $n = 2$, chaque tirage donne nécessairement i ou j : la variable vaut constamment k , ce qui correspond bien à la loi binomiale de paramètre 1 et à une variance nulle.

- b. Méthode 1 — Variance d'une somme

La formule $V(X_i + X_j) = V(X_i) + V(X_j) + 2\text{Cov}(X_i, X_j)$ est valable sans supposer l'indépendance. Les résultats précédents donnent :

$$2 \operatorname{Cov}(X_i, X_j) = \frac{2k(n-2)}{n^2} - \frac{2k(n-1)}{n^2} = -\frac{2k}{n^2}.$$

Ainsi $\operatorname{Cov}(X_i, X_j) = -k/n^2$.

Méthode 2 — Indicatrices des tirages

Pour $r \in \llbracket 1; k \rrbracket$, notons $I_{r,i}$ l'indicatrice de l'obtention du numéro i au tirage r . Alors

$$X_i = \sum_{r=1}^k I_{r,i}.$$

Si $r \neq s$, les tirages indépendants rendent $I_{r,i}$ et $I_{s,j}$ indépendantes, donc leur covariance est nulle. Pour un même tirage, $i \neq j$ donne $I_{r,i}I_{r,j} = 0$, tandis que leurs espérances valent $1/n$. Donc :

$$\operatorname{Cov}(I_{r,i}, I_{r,j}) = E(I_{r,i}I_{r,j}) - E(I_{r,i})E(I_{r,j}) = -\frac{1}{n^2}.$$

La bilinéarité de la covariance, appliquée à des sommes finies, conduit à

$$\operatorname{Cov}(X_i, X_j) = \sum_{r=1}^k (-1/n^2) = -k/n^2.$$

Cette approche distingue la dépendance négative entre numéros au sein d'un tirage de l'indépendance entre tirages.

Partie II

1. Un premier tirage fournit toujours un seul numéro : $P(Z_1 = 1) = 1$ et $E(Z_1) = 1$.

Pour deux tirages, $Z_2 = 1$ si le deuxième numéro est égal au premier, ce qui a probabilité $1/n$. Sinon, $Z_2 = 2$. Ainsi :

$$P(Z_2 = 1) = \frac{1}{n}, \quad P(Z_2 = 2) = \frac{n-1}{n},$$

$$E(Z_2) = \frac{1}{n} + 2 \frac{n-1}{n} = 2 - \frac{1}{n}.$$

2. On fixe un entier $k \geq 1$.

a. Avoir $Z_k = 1$ signifie que tous les tirages portent le même numéro. Pour chacun des n numéros, la probabilité de cette suite constante est n^{-k} ; ces événements sont disjoints. Donc :

$$P(Z_k = 1) = n^{1-k}.$$

Si $1 \leq k \leq n$, l'égalité $Z_k = k$ signifie que les k tirages sont tous distincts. Le premier numéro est quelconque, le second doit éviter le premier, le troisième les deux précédents, et ainsi de suite. Les probabilités composées donnent :

$$P(Z_k = k) = 1 \times \frac{n-1}{n} \times \dots \times \frac{n-k+1}{n} = \frac{n!}{(n-k)! n^k}.$$

Pour $k = 1$, ce produit vaut 1. Si $k > n$, plus de n numéros distincts sont impossibles, d'où $P(Z_k = k) = 0$.

- b.** Posons $p_{k,j} = P(Z_k = j)$, avec en particulier $p_{k,0} = 0$. Pour atteindre ℓ numéros distincts après le tirage suivant, il existe deux cas disjoints : on en avait déjà ℓ et le dernier tirage donne un numéro déjà vu, ou on en avait $\ell - 1$ et le dernier tirage donne un numéro nouveau.

Pour toute suite fixée des k premiers tirages comportant j numéros distincts, le tirage indépendant suivant appartient à ces j numéros avec probabilité j/n ; il appartient à leur complément avec probabilité $(n - j)/n$. En sommant sur ces suites, toutes de probabilité positive, on obtient :

$$p_{k+1,\ell} = \frac{\ell}{n} p_{k,\ell} + \frac{n - \ell + 1}{n} p_{k,\ell-1}, \quad 1 \leq \ell \leq n.$$

Ce raisonnement reste valable lorsque l'une des valeurs de Z_k est impossible : sa contribution est alors simplement nulle. Il n'utilise aucune probabilité conditionnelle sachant un événement impossible.

- c.** Puisque $1 \leq Z_k \leq n$, toutes les sommes suivantes sont finies. En multipliant II.2.b par ℓ , puis en sommant de 1 à n :

$$E(Z_{k+1}) = \frac{1}{n} \sum_{\ell=1}^n \ell^2 p_{k,\ell} + \frac{1}{n} \sum_{\ell=1}^n \ell(n - \ell + 1) p_{k,\ell-1}.$$

Dans la seconde somme, posons $j = \ell - 1$. Le terme $j = 0$ est nul car $p_{k,0} = 0$; on peut ajouter le terme $j = n$ car son coefficient $n - j$ est nul. Ainsi :

$$E(Z_{k+1}) = \frac{1}{n} \sum_{j=1}^n (j^2 + (j+1)(n-j)) p_{k,j}.$$

Or $j^2 + (j+1)(n-j) = (n-1)j + n$. Comme

$$\sum_{j=1}^n p_{k,j} = 1,$$

il vient :

$$E(Z_{k+1}) = \frac{n-1}{n}E(Z_k) + 1.$$

3. Méthode 1 — Suite arithmético-géométrique

Avec $v_k = E(Z_k) - n$, la relation précédente devient :

$$v_{k+1} = \frac{n-1}{n}E(Z_k) + 1 - n = \frac{n-1}{n}v_k.$$

La suite (v_k) est donc géométrique de raison $q = (n-1)/n$. Son premier terme vaut $v_1 = 1 - n = -nq$. Pour tout $k \geq 1$, $v_k = v_1 q^{k-1} = -nq^k$, d'où :

$$E(Z_k) = n \left(1 - \left(\frac{n-1}{n} \right)^k \right).$$

Méthode 2 — Indicatrices des numéros présents

Pour chaque $i \in \llbracket 1; n \rrbracket$, notons J_i l'indicatrice de la présence du numéro i parmi les k premiers tirages. Alors

$$Z_k = \sum_{i=1}^n J_i.$$

La probabilité de ne jamais obtenir i vaut $((n-1)/n)^k$, par indépendance des tirages. Donc :

$$E(J_i) = P(J_i = 1) = 1 - \left(\frac{n-1}{n} \right)^k.$$

La linéarité de l'espérance donne directement la même formule pour $E(Z_k)$, sans supposer que les indicatrices J_i sont indépendantes. Cette méthode fournit une vérification directe du résultat, tandis que la première exploite la relation de récurrence demandée.

Partie III

1. En prenant $n = 4$ dans II.2.a :

$$P(Z_k = 1) = 4^{1-k} = \frac{4}{4^k}.$$

Il n'y a que quatre numéros dans l'urne, donc $Z_k \leq 4$ et $P(Z_k \geq 5) = 0$.

2. Parmi les quatre numéros, une paire non ordonnée se choisit de $4 \times 3/2 = 6$ façons : chaque choix ordonné de deux numéros distincts compte deux fois la même paire.

Fixons une telle paire. Les suites de longueur k ne contenant que ses deux numéros sont au nombre de 2^k , car chacun des k emplacements offre deux choix indépendants. Parmi elles, deux sont constantes ; les $2^k - 2$ autres utilisent bien les deux numéros. Chacune des suites possibles de tirages dans l'urne a probabilité 4^{-k} .

Les six événements correspondant aux six paires effectivement présentes sont disjoints. En additionnant leurs probabilités :

$$P(Z_k = 2) = 6 \frac{2^k - 2}{4^k}.$$

3. On utilise les événements A_i définis dans l'énoncé : le numéro i est absent des k premiers tirages.

a. L'événement $[Z_k \leq 3]$ signifie qu'au moins un numéro manque, donc qu'au moins l'un des événements A_1, A_2, A_3, A_4 est réalisé.

On établit la formule particulière demandée sans supposer une formule générale de réunion. Pour une issue fixée, notons r le nombre de numéros absents. Puisqu'au moins un tirage a lieu, $r \in \{0, 1, 2, 3\}$. Comptons combien de fois cette issue est comptée dans la somme des indicatrices des A_i , moins celle des intersections par paires, plus celle des intersections par triplets.

Si $r = 0$, sa contribution vaut 0. Si $r = 1$, elle vaut 1. Si $r = 2$, elle vaut $2 - 1 = 1$. Si $r = 3$, elle vaut $3 - 3 + 1 = 1$. Dans tous les cas, on obtient exactement l'indicatrice de $[Z_k \leq 3]$. Par linéarité de l'espérance :

$$P(Z_k \leq 3) = \sum_{i=1}^4 P(A_i) - \sum_{1 \leq i < j \leq 4} P(A_i \cap A_j) + \sum_{1 \leq i < j < h \leq 4} P(A_i \cap A_j \cap A_h).$$

Par symétrie des quatre numéros, toutes les probabilités simples sont égales, toutes les probabilités d'intersections de deux événements sont égales et il en va de même pour trois événements. Il y a quatre termes simples, six paires et quatre triplets. Par conséquent :

$$P(Z_k \leq 3) = 4P(A_1) - 6P(A_1 \cap A_2) + 4P(A_1 \cap A_2 \cap A_3).$$

b. Ne pas tirer la boule 1 impose de choisir à chaque tirage l'un des trois autres numéros. L'indépendance donne $P(A_1) = (3/4)^k$.

Pour ne tirer ni 1 ni 2, il reste deux numéros : $P(A_1 \cap A_2) = (2/4)^k = (1/2)^k$.

Pour ne tirer aucun des numéros 1, 2, 3, tous les tirages doivent donner 4 :

$$P(A_1 \cap A_2 \cap A_3) = (1/4)^k.$$

c. En remplaçant dans III.3.a :

$$P(Z_k \leq 3) = 4 \left(\frac{3}{4}\right)^k - 6 \left(\frac{1}{2}\right)^k + 4 \left(\frac{1}{4}\right)^k = \frac{4 \cdot 3^k - 6 \cdot 2^k + 4}{4^k}.$$

Les valeurs possibles de Z_k sont 1, 2, 3, 4. On retrace donc les probabilités déjà calculées :

$$P(Z_k = 3) = P(Z_k \leq 3) - P(Z_k = 1) - P(Z_k = 2)$$

$$P(Z_k = 3) = \frac{4 \cdot 3^k - 6 \cdot 2^k + 4 - 4 - 6(2^k - 2)}{4^k} = \frac{4 \cdot 3^k - 12 \cdot 2^k + 12}{4^k}.$$

Enfin, par passage au complémentaire :

$$P(Z_k = 4) = 1 - P(Z_k \leq 3) = \frac{4^k - 4 \cdot 3^k + 6 \cdot 2^k - 4}{4^k}.$$

Ces expressions, avec celles de III.1 et III.2, donnent la loi complète de Z_k . Leur somme vaut 1, conformément à la partition par les quatre valeurs possibles.

Relevé des adaptations

Exercice 1, II.4.c : code adapté en Python. La demande de rédaction dans l'ancien langage a été remplacée par une demande identique en Python ; les commandes utiles de NumPy et l'affichage sont documentés. L'objectif de dichotomie et la précision 10^{-3} sont conservés.