

AP 4

## Trace d'une matrice carrée - Indice de Gini

### Exercice 1 (Trace d'une matrice carrée)

Soit  $E = \mathcal{M}_n(\mathbb{R})$  et  $t$  l'application de  $E$  dans  $\mathbb{R}$  définie par :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \quad t(A) = \sum_{i=1}^n [A]_{i,i}.$$

On dit que  $t(A)$  est la trace de la matrice  $A$ .

1. Calculer  $t(I_n)$ .
2. Pour les 5/2.
  - (a) Montrer que  $t$  est linéaire.
  - (b) Calculer l'image de  $t$ , et en donner une base.
  - (c) En déduire  $\dim(\text{Ker}(t))$ , puis déterminer une base de  $\text{Ker}(t)$ .
3.
  - (a) Montrer que pour tout  $A \in E$ ,  $t({}^t A) = t(A)$ .
  - (b) Montrer que, pour tout  $A, B \in \mathcal{M}_n(\mathbb{R})$ ,  $t(AB) = t(BA)$ .
  - (c) En déduire que, si  $A = PBP^{-1}$  avec  $P$  une matrice inversible, alors  $t(A) = t(B)$ .
  - (d) Soit  $A$  une matrice de  $\mathcal{M}_n(\mathbb{R})$  diagonalisable admettant  $\lambda_1 < \lambda_2 < \dots < \lambda_p$  comme valeurs propres distinctes. Exprimer  $t(A)$  en fonction des valeurs propres et des dimensions des sous-espaces propres de  $A$ .

### Exercice 2 (Indice de Gini)

Étudier l'évolution des inégalités dans la répartition des richesses, matérielles ou symboliques, dans une société est un des thèmes majeurs des sciences humaines. Considérons un exemple élémentaire. Le tableau ci-dessous présente le pourcentage d'accès à l'enseignement secondaire en Grande-Bretagne lors de deux périodes pour deux catégories sociales :

	avant 1910	entre 1935 et 1940
Profession libérale	37%	62%
Ouvriers	1%	10%

On propose trois modes de comparaison des inégalités entre les deux classes sociales.

1. En regardant l'augmentation des pourcentages pour les deux classes entre les deux périodes, on conclut que l'inégalité a augmenté entre la classe la plus aisée (Profession libérale) et la plus défavorisée (Ouvriers).
2. Si on regarde le taux d'accroissement des pourcentages, comme  $\frac{10}{1} > \frac{62}{37}$ , on déduit que l'inégalité a diminué.
3. Si on regarde le taux d'accroissement des pourcentages *de ceux qui n'accèdent pas à l'enseignement secondaire*, comme  $\frac{90}{99} > \frac{38}{63}$ , on déduit que l'inégalité a augmenté puisque le nombre de ceux qui n'ont pas accès à l'enseignement supérieur a proportionnellement plus diminué que celui de ceux qui y ont accès.

Comme on le voit chacune des façons de voir est légitime à sa manière. L'objet du problème est d'introduire des outils afin d'étudier la *concentration* d'une loi de probabilité pour contourner des paradoxes auxquels une analyse trop rapide peut conduire, ou du moins d'en être conscient.

## 1 Indice de Gini

On rappelle qu'une fonction numérique définie sur l'intervalle  $J$  de  $\mathbb{R}$  est *convexe* sur  $J$  si elle vérifie la propriété suivante :  $\forall (t_1, t_2) \in J^2, \forall \lambda \in [0, 1], f(\lambda t_1 + (1 - \lambda)t_2) \leq \lambda f(t_1) + (1 - \lambda)f(t_2)$ .

On rappelle en outre qu'une fonction  $f$  est *concave* si  $-f$  est convexe.

On désigne par  $E$  l'ensemble des applications  $f$  définies sur  $[0, 1]$  à valeurs dans  $[0, 1]$ , continues et convexes sur  $[0, 1]$ , et telles que  $f(0) = 0$  et  $f(1) = 1$ . Pour toute application  $f$  de  $E$ , on note  $\tilde{f}$  l'application associée à  $f$ , définie sur  $[0, 1]$  par  $\tilde{f}(t) = t - f(t)$ .

On pose enfin  $I(f) = 2 \int_0^1 \tilde{f}(t)dt = 2 \int_0^1 (t - f(t))dt$ .  $I(f)$  s'appelle l'**indice de Gini** de l'application  $f$ .

1. (a) Donnez une interprétation géométrique de la propriété de convexité.  
 (b) Lorsque  $f$  est une fonction de classe  $C^1$  sur  $[0, 1]$ , rappeler la caractérisation de la convexité de  $f$  sur  $[0, 1]$  à l'aide de la dérivée  $f'$ .
2. (a) Justifier que  $\tilde{f}$  est concave sur  $[0, 1]$ .  
 (b) Montrer que  $I(f) = 1 - 2 \int_0^1 f(t)dt$ .  
 (c) Représenter dans un même repère orthonormé les fonctions  $f$  et  $t \mapsto t$  et donner une interprétation géométrique de  $I(f)$ .

### 3. Un premier exemple.

Soit  $f : [0, 1] \rightarrow \mathbb{R}$  telle que  $f(t) = t^2$  pour tout  $t \in [0, 1]$ .

- (a) Montrer que  $f$  est un élément de  $E$ .
- (b) Calculer  $I(f)$ .

### 4. Propriétés de l'indice de Gini.

- (a) Pour  $f$  élément de  $E$ , établir que  $I(f) \geq 0$ .
- (b) Montrer que  $I(f) = 0$  si et seulement si  $f(t) = t$  pour tout  $t \in [0, 1]$ .
- (c) Montrer que pour tout  $f$  élément de  $E$ ,  $I(f) < 1$ .
- (d) Pour tout entier  $n > 0$ , on définit  $f_n$  sur  $[0, 1]$  par  $f_n(t) = t^n$ .
  - i. Pour tout entier  $n$  strictement positif, calculer  $I(f_n)$ .
  - ii. En déduire que pour tout réel  $A$  vérifiant  $0 \leq A < 1$ , il existe  $f$  appartenant à  $E$  telle que  $I(f) > A$ .

### 5. Minoration de l'indice de Gini

- (a) Soit  $f$  élément de  $E$ . Montrer qu'il existe  $t_0$  dans  $]0, 1[$  tel que  $\tilde{f}(t_0) = \max_{t \in [0, 1]} \tilde{f}(t)$ .
- (b) Montrer que pour tout  $t$  de  $[0, t_0]$ ,  $\tilde{f}(t) \geq \tilde{f}(t_0) \cdot \frac{t}{t_0}$ .
- (c) Montrer que pour tout  $t$  de  $[t_0, 1]$ ,  $\tilde{f}(t) \geq \tilde{f}(t_0) \cdot \frac{t-1}{t_0-1}$ .
- (d) En déduire que  $I(f) \geq \tilde{f}(t_0)$ .

L'indice de Gini donne une indication sur la concentration des richesses d'un pays si l'on suppose que la fonction  $f$  rend compte de cette concentration. Par exemple,  $f(0, 3) = 0, 09$  s'interprète par le fait que dans la population classée par ordre de richesse croissante, les premiers 30% de la population possèdent 9% de la richesse totale du pays. Plus l'indice  $I(f)$  est grand, plus la répartition des richesses est inégalitaire.

## 2 Application à une population

Une population de  $N$  personnes est divisée en deux classes (typiquement hommes et femmes) et en  $n$  catégories (par exemple socio-professionnelles), suivant le tableau à double entrée suivant où tous les  $x_i$  et  $y_i$  pour  $i$  dans  $\llbracket 1, n \rrbracket$  sont des entiers naturels.

On suppose en outre que pour tout  $i$  dans  $\llbracket 1, n \rrbracket$ ,  $x_i \neq 0$ .

Classes \ Catégories	$c_1$	$c_2$	$c_3$	...	$c_i$	...	$c_n$	Total
I	$x_1$	$x_2$	$x_3$	...	$x_i$	...	$x_n$	$X$
II	$y_1$	$y_2$	$y_3$	...	$y_i$	...	$y_n$	$Y$
Total	$n_1$	$n_2$	$n_3$	...	$n_i$	...	$n_n$	$N$

où on a donc posé  $X = \sum_{i=1}^n x_i$ ,  $Y = \sum_{i=1}^n y_i$  et  $X + Y = N$ . On suppose en outre que  $Y > 0$ .

Pour  $i$  appartenant à  $\llbracket 1, n \rrbracket$ , on adopte les notations suivantes :

$$p_i = \frac{n_i}{N}, \quad q_i = \frac{x_i}{X}, \quad r_i = \frac{y_i}{Y}.$$

On note aussi  $\varepsilon_i = \frac{x_i}{n_i}$ , et  $\varepsilon = \frac{X}{N}$ , et on suppose que les catégories sont numérotées de telle sorte que

$$\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_n.$$

6. On pose  $\Omega = \{c_1, c_2, \dots, c_n\}$ , ensemble des catégories dans la population.

(a) Montrer que  $P = (p_i)_{1 \leq i \leq n}$ ,  $Q = (q_i)_{1 \leq i \leq n}$  et  $R = (r_i)_{1 \leq i \leq n}$  sont des distributions de probabilités.

(b) Montrer que  $\frac{q_1}{p_1} \leq \dots \leq \frac{q_n}{p_n}$  (\*)

(c) Montrer que  $\frac{r_1}{p_1} \geq \dots \geq \frac{r_n}{p_n}$

(d) Montrer que pour  $i$  appartenant à  $\llbracket 1, n \rrbracket$ ,  $r_i = \frac{n_i - x_i}{N - X} = \frac{p_i - \varepsilon q_i}{1 - \varepsilon}$ .

7. Dans un premier temps, nous allons construire une application appartenant à  $E$ , qui permet de mesurer les inégalités à l'intérieur de la classe I.

On pose  $P_0 = Q_0 = 0$ , et pour  $i \in \llbracket 1, n \rrbracket$ ,  $P_i = \sum_{h=1}^i p_h$  et  $Q_i = \sum_{h=1}^i q_h$ . On définit alors

l'application  $\varphi$  de  $[0, 1]$  dans  $[0, 1]$  telle que, pour tout entier  $i \in \llbracket 0, n \rrbracket$ ,  $\varphi(P_i) = Q_i$  et pour tout entier  $i \in \llbracket 0, n - 1 \rrbracket$ ,  $\varphi$  est affine sur le segment  $[P_i, P_{i+1}]$ .

(a) On suppose **dans cette question**  $n = 3$ .

Représenter dans un repère orthonormé  $\varphi$  lorsque  $P = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$  et  $Q = \left(\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\right)$ .

(b) Montrer que, dans le plan rapporté à un repère orthonormé, la pente de la droite passant par les points de coordonnées  $(P_{i-1}, Q_{i-1})$  et  $(P_i, Q_i)$  est  $u_i = \frac{q_i}{p_i}$  pour  $i$  appartenant à  $\llbracket 1, n \rrbracket$ .

(c) Montrer que si  $i \in \llbracket 0, n - 1 \rrbracket$  et  $t \in [P_i, P_{i+1}]$ , on a  $\varphi(t) = u_{i+1}(t - P_i) + Q_i$ .

(d) **En admettant** que les inégalités (\*) de la question 6.(b) permettent d'affirmer que  $\varphi$  est convexe, justifier que  $\varphi$  appartient à  $E$ .

(e) Pour  $i \in \llbracket 0, n - 1 \rrbracket$ , calculer  $\int_{P_i}^{P_{i+1}} \varphi(t) dt$ .

(f) Exprimer  $I(\varphi)$  sous la forme d'une somme en fonction de  $P_0, P_1, \dots, P_n, Q_0, \dots, Q_n$ .

8. Nous allons maintenant étudier l'application correspondante pour la classe II.

On pose  $P_0 = R_0 = 0$  et pour  $i \in \llbracket 1, n \rrbracket$ ,  $P_i = \sum_{h=1}^i p_h$  et  $R_i = \sum_{h=1}^i r_h$ . De même, on définit pour  $i$  élément de  $\llbracket 0, n \rrbracket$ ,  $\Pi_i = 1 - P_{n-i}$ . On considère l'application  $\psi$  de  $[0, 1]$  dans  $[0, 1]$  telle que pour tout  $i \in \llbracket 0, n \rrbracket$ ,  $\psi(P_i) = R_i$  et pour tout entier  $i \in \llbracket 0, n-1 \rrbracket$ ,  $\psi$  est affine sur le segment  $[P_i, P_{i+1}]$ .

(a) Montrer que la pente de la droite passant par les points de coordonnées  $(P_{i-1}, R_{i-1})$  et  $(P_i, R_i)$  est  $v_i = \frac{r_i}{p_i}$  pour  $i \in \llbracket 1, n \rrbracket$ .

(b) On considère l'application  $\psi^*$  définie pour tout  $t \in [0, 1]$ , par  $\psi^*(t) = 1 - \psi(1 - t)$ .

i. On suppose **dans cette question**  $n = 3$ .

Représenter dans un même repère orthonormé les courbes représentatives de  $\psi$  et  $\psi^*$  lorsque  $P = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$  et  $R = \left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right)$ .

ii. Montrer que  $\psi^*$  est convexe sur  $[0, 1]$ .

iii. Montrer que  $\psi^*$  est affine sur  $[\Pi_{i-1}, \Pi_i]$  pour  $i \in \llbracket 0, n-1 \rrbracket$ .

iv. Montrer que la pente de  $\psi^*$  sur  $[\Pi_{i-1}, \Pi_i]$  est  $v_{n-i+1}$  pour  $i \in \llbracket 1, n \rrbracket$ .

On dit dans cette situation que les fonctions  $\varphi$  et  $\psi^*$  sont **adjointes** l'une de l'autre. C'est leur comparaison que Gini a proposé de considérer pour "mesurer les inégalités" entre la population de catégorie I et celle de catégorie II.

Une égalité entre les fonctions adjointes signale notamment l'absence totale d'inégalité sociale. La dernière question précise quelque peu ce point.

9. (a) Montrer que si  $\varphi = \psi^*$  alors pour tout  $i$  appartenant à  $\llbracket 1, n \rrbracket$ ,

$$\frac{\varepsilon_i}{\varepsilon} = \frac{1 - \varepsilon_{n-i+1}}{1 - \varepsilon}.$$

(b) Montrer que si  $\varphi = \psi^*$ , alors pour tout  $i$  appartenant à  $\llbracket 1, n \rrbracket$ ,  $\varepsilon_i + \varepsilon_{n-i+1} = 2\varepsilon$ .

(c) Dédurre que si  $\varphi = \psi^*$ , on a pour tout  $i$  appartenant à  $\llbracket 1, n \rrbracket$ ,  $\varepsilon_i(1 - 2\varepsilon) = \varepsilon(1 - 2\varepsilon)$

(d) On suppose que  $\varepsilon \neq \frac{1}{2}$ . Montrer que si  $\varphi = \psi^*$ , alors pour tout  $i$  appartenant à  $\llbracket 1, n \rrbracket$ ,  $\varepsilon_i = \varepsilon$ . Interpréter ce résultat.