MATHÉMATIQUES ET SCIENCES HUMAINES

ANTOINE VALEYRE

Formes et propriétés des indices d'inégalité entre proportions

Mathématiques et sciences humaines, tome 132 (1995), p. 13-37

<http://www.numdam.org/item?id=MSH_1995__132__13_0>

Math. Inf. Sci. hum. (33e année, n°132, *1995*, pp.13-37)

FORMES ET PROPRIÉTÉS DES INDICES D'INÉGALITÉ ENTRE PROPORTIONS

Antoine VALEYRE[1]

RÉSUMÉ — *Cet article porte sur les indicateurs qui permettent de comparer les inégalités de proportions entre deux catégories. Après avoir précisé les propriétés qui caractérisent la définition d'indices d'inégalité entre deux proportions, il analyse celles qu'il est souhaitable de leur appliquer, notamment les propriétés de cohérence ou d'homogénéité dont il montre l'incompatibilité. Il examine différents modes de construction d'indices d'inégalité : à partir de mesures, de distances, de caractéristiques statistiques de dispersion, des indices d'Atkinson et de l'indice de concentration de Gini. Il étudie les propriétés et les formes de ces indices, ainsi que les relations qui s'établissent entre eux. Enfin il traite des comparaisons qualitatives d'inégalité qui se fondent sur les fonctions de concentration de Gini.*

SUMMARY — Forms and properties of inequality indexes between proportions
This article deals with the indexes which allow to compare the inequalities of proportions between two classes. After examining the properties which characterize the définition of inequality indexes between two proportions, it analyzes some other properties like coherence or homogeneity whose incompatibility is proved. It is concerned by different kinds of inequality indexes which are built on measures, distances, statistical measures of dispersion, Atkinson's indexes and Gini's coefficient. It studies the forms and the properties of these indexes and it shows the relationships between them. Finally it deals with the qualitatives comparisons of inequalities based on the Gini's concentration functions.

INTRODUCTION

Les méthodes d'évaluation et de comparaison des inégalités entre proportions ont été l'objet d'un débat substantiel dans la *Revue Française de Sociologie* et la revue *Mathématiques et Sciences humaines* avec la publication de nombreux articles sur ce thème il y a une dizaine d'années[2]. Ce débat engagé par des sociologues de l'éducation trouve son origine dans les divergences d'appréciation sur l'évolution des inégalités sociales de scolarisation, auxquelles conduisent les différents indicateurs d'usage courant. Il s'est plus particulièrement développé

[1] Laboratoire Techniques, Territoires et Sociétés, Central IV, 1 avenue Montaigne, 93167 Noisy-le-Grand cedex, et Centre d'Etudes de l'Emploi, Le Descartes I, 29 promenade Michel Simon, 93166 Noisy-le-Grand cedex.

[2] Ont successivement contribué à ce débat dans la *Revue Française de Sociologie* en 1984 et 1985, les articles de J.-C. Combessie [8, 9], J.-P. Florens [10], J.-P. Grémy [12], M. Barbut [3], J. Prévot [15] et D. Merllié [14]. Se rattachent également à ce débat des articles de M. Barbut [4, 5] et de J.-P. Grémy [13] publiés à la même époque dans *Mathématiques et Sciences humaines* et un article de L.-A. Vallet paru plus tardivement dans la *Revue Française de Sociologie* [20].

autour de l'exemple des inégalités d'accès à l'enseignement secondaire anglais pendant la première moitié du siècle entre deux catégories sociales contrastées, les enfants de cadres et les enfants d'ouvriers, à partir des travaux de John Westergaard et Alan Little [21][3] et de Raymond Boudon [6].

Cet article présente les résultats d'une étude qui s'inscrit dans le prolongement de ce débat [19][4]. Il porte sur les indicateurs permettant de comparer les inégalités de proportions entre deux catégories, par exemple les inégalités de taux d'admission dans l'enseignement secondaire entre enfants de cadres et enfants d'ouvriers.

Une première partie présente les variables, proportions ou distributions, qui caractérisent les situations d'inégalité et qui constituent les éléments sur lesquels portent les indicateurs d'inégalité. Une seconde partie précise les propriétés retenues pour définir des indices d'inégalité entre proportions et les propriétés qu'il est souhaitable de leur appliquer. Les parties suivantes étudient différents types d'indices d'inégalité et analysent leurs formes et leurs propriétés. Elles examinent successivement différents modes de construction d'indices d'inégalité, à partir de mesures, de distances entre proportions ou entre couples de proportions dichotomiques, de caractéristiques statistiques de dispersion, ou des indices d'Atkinson. Enfin une dernière partie est consacrée aux comparaisons d'inégalité entre distributions dichotomiques à l'aide des fonctions de concentration de Gini qui permettent dans certains cas de se dispenser d'indicateurs quantitatifs.

1. VARIABLES REPRÉSENTATIVES DES SITUATIONS D'INÉGALITÉ

Pour apprécier l'évolution de l'inégalité entre les proportions de deux catégories[5], par exemple entre les taux de scolarisation des enfants de cadres et d'ouvriers, de nombreuses méthodes peuvent être adoptées. Toutes ne sont pas présentées ici, notamment celles qui prennent en compte les taux de croissance des proportions. Cet article concerne exclusivement les méthodes consistant à comparer des indicateurs quantitatifs ou qualitatifs caractérisant les situations d'inégalité à chaque date de la période de référence. Ce type de méthodes soulève alors la question de la détermination des variables représentatives des situations d'inégalité à une date donnée et de facon liée la question de la définition d'indicateurs portant sur ces variables.

1.1. Tableau statistique de référence

Les variables représentatives des situations d'inégalité entre deux catégories à une date donnée, par exemple l'inégalité de scolarisation entre enfants de cadres et d'ouvriers, sont issues des données d'effectifs de tableaux statistiques à double entrée croisant la variable "situation scolaire" comportant les modalités "admis" et "exclus" et la variable "catégorie sociale du père" comportant les modalités "cadre" et "ouvrier". Ces tableaux sont de la forme :

		Situation scolaire		
		Admis	Exclus	Ensemble
Catégorie sociale du père	Cadre	N_{11}	N_{12}	$N_{1\bullet}$
	Ouvrier	N_{21}	N_{22}	$N_{2\bullet}$
	Total	$N_{\bullet 1}$	$N_{\bullet 2}$	$N_{\bullet\bullet}$

Tableau T1
Tableau statistique de référence

[3] Les références entre crochets renvoient à la bibliographie.

[4] L'auteur remercie un rapporteur anonyme pour les nombreuses remarques et suggestions qui ont permis d'améliorer cette présentation.

[5] L'analyse présentée ici porte sur des comparaisons temporelles. Elle vaut bien évidemment pour toute autre forme de comparaison, notamment pour les comparaisons spatiales.

1.2. Représentation des situations d'inégalité par des proportions

La démarche adoptée par de nombreux sociologues de l'éducation pour définir des indicateurs quantitatifs d'inégalité consiste à représenter les situations d'inégalité sociale de scolarisation par le couple $(x,y)=(N_{11}/N_{1\bullet},N_{21}/N_{2\bullet})$ des proportions d'admis parmi les enfants de cadres et parmi les enfants d'ouvriers, donc par le couple des fréquences d'admission conditionnelles à la catégorie cadre et à la catégorie ouvrier. La représentation des situations d'inégalité peut également se fonder sur le couple $(x^*,y^*)=(N_{12}/N_{1\bullet},N_{22}/N_{2\bullet})$ des proportions d'exclus, complémentaires des proportions d'admis ($x^*=1-x$ et $y^*=1-y$).

Des variables supplémentaires sont parfois adjointes au couple de proportions d'admis (ou d'exclus) pour décrire plus complètement les situations d'inégalité. Ainsi les indicateurs d'inégalité relatifs introduisent dans leur définition une proportion de référence caractéristique de chaque situation d'inégalité, afin de relativiser la prise en compte des deux proportions qu'ils confrontent. Comme proportion de référence on peut choisir par exemple la proportion $m=N_{\bullet1}/N_{\bullet\bullet}$ d'admis parmi l'ensemble des enfants de cadres et d'ouvriers (ou la proportion d'exclus $m^*=1-m=N_{\bullet2}/N_{\bullet\bullet}$), c'est à dire la fréquence marginale d'admission (ou d'exclusion). Si les données sont disponibles, on peut également adopter la proportion t d'admis parmi les enfants de l'ensemble de la population active (ou la proportion d'exclus $t^*=1-t$).

De même les indicateurs d'inégalité pondérés prennent en compte, outre le couple de proportions des deux catégories sociales concernées, la distribution marginale de leurs fréquences $(a,a^*)=(N_{1\bullet}/N_{\bullet\bullet},N_{2\bullet}/N_{\bullet\bullet})$.

1.3. Représentation des situations d'inégalité par des distributions

Le choix des couples de proportions d'admis (ou d'exclus) comme variables de base pour construire des indicateurs d'inégalité n'est pas la seule démarche possible. Les variables représentatives des situations d'inégalité peuvent aussi être constituées par des couples de distributions dichotomiques de situations scolaires ou de catégories sociales établies sur la base du tableau statistique de référence.

Ainsi on peut construire des indicateurs fondés sur l'un des couples de distributions conditionnelles et marginale des catégories sociales. Ces couples combinent deux à deux les distributions des catégories sociales conditionnelle à l'admission $(u,u^*)=(N_{11}/N_{\bullet1},N_{21}/N_{\bullet1})=(ax/m,a^*y/m)$, conditionnelle à l'exclusion $(v,v^*)=(N_{12}/N_{\bullet2},N_{22}/N_{\bullet2})=(ax^*/m^*,a^*y^*/m^*)$ ou marginale (a,a^*), c'est-à-dire les distributions d'enfants de cadres et d'ouvriers que l'on observe parmi les admis, parmi les exclus ou dans l'ensemble.

On peut également construire des indicateurs sur la base de l'un des couples de distributions conditionnelles ou marginale des situations scolaires. Les variables représentatives des situations d'inégalité sont alors des couples de distributions de proportions dichotomiques d'admis et d'exclus. On peut choisir par exemple le couple des distributions de proportions d'admis et d'exclus observées parmi les enfants de cadres $(x,x^*)=(N_{11}/N_{1\bullet},N_{12}/N_{1\bullet})$ et parmi les enfants d'ouvriers $(y,y^*)=(N_{21}/N_{2\bullet},N_{22}/N_{2\bullet})$, c'est à dire les couples des distributions des situations scolaires conditionnelles à la catégorie cadre et à la catégorie ouvrier.

1.4. Les données d'inégalité sociale de scolarisation utilisées

L'exemple retenu à titre d'illustration porte sur l'accès à l'enseignement secondaire britannique des enfants d'ouvriers et de cadres[6] au cours de la première moitié du siècle. Il se fonde sur les données du rapport de John Westergaard et Alan Little auxquelles se réfèrent de nombreux articles s'inscrivant dans le débat sur les comparaisons d'inégalité entre proportions. Les données de proportions d'admis et de fréquences utilisées dans cet article sont récapitulées dans le tableau T2[7]. Les données de proportions d'exclus et plus généralement des différentes distributions envisagées pour représenter les situations d'inégalité sociale de scolarisation s'en déduisent.

Catégorie sociale du père	Taux d'admis (en %)		Fréquence (en %)			
	Naissance <1910	Naissance 1935-40	Naissance < 1910		Naissance 1935-40	
Cadres	37	62	19	*46*	14	*31*
Ouvriers	1	10	22	*54*	31	*69*
Cadres et ouvriers	18	26	41	*100*	45	*100*
Ensemble des catégories	12	23	100	-	100	-

Tableau T2
Accès à l'enseignement secondaire britannique selon les catégories sociales

2. DÉFINITION ET PROPRIÉTÉS DES INDICES D'INÉGALITÉ ENTRE PROPORTIONS

2.1. Propriétés de définition des indices d'inégalité entre proportions

Par définition, un indice d'inégalité entre deux proportions est une fonction M définie sur un domaine carré I^2 ($I=[0,1]$, $]0,1]$, $[0,1[$ ou $]0,1[$) qui respecte les propriétés suivantes :

Positivité : $\forall (x,y)\in I^2 \quad M(x,y)\geq 0$;
Identité : $m_0=\mathrm{Min}_{(x,y)\in I^2}M(x,y) \quad \forall (x,y)\in I^2 \quad M(x,y)=m_0 \Leftrightarrow x=y$;
Symétrie : $\forall (x,y)\in I^2 \quad M(y,x)=M(x,y)$;
Variation en V : $\forall (x,y)\in I^2$ tel que $x\neq y \ \forall z\in]x,y[\quad M(x,z)$ et $M(z,y) < M(x,y)$.

La propriété d'identité caractérise les indices qui prennent leur valeur minimum lorsque les proportions sont égales et seulement dans ce cas. Elle est introduite pour ne donner leur valeur minimum aux indices que dans les situations d'absence d'inégalité entre proportions.

La propriété de symétrie caractérise l'invariance des indices avec l'ordre des deux proportions x et y. Elle conduit à attribuer la même valeur à l'inégalité de y par rapport à x et à celle de x par rapport à y. Elle est proposée pour éviter les incohérences d'interprétation dans les comparaisons [12, p.410].

---

[6] Plus précisément, l'inégalité sociale est étudiée entre deux catégories extrêmes, la classe supérieure regroupant les dirigeants, les professions libérales et les cadres supérieurs et moyens, et la catégorie des ouvriers semi ou non qualifiés. Pour simplifier on utilisera les désignations de cadres et d'ouvriers pour caractériser ces deux catégories contrastées. L'importance relative de la catégorie des cadres en début de période s'explique probablement par la proportion élevée de petits patrons qu'elle comporte, à une époque où le mouvement de concentration économique était encore peu prononcé.

[7] Les fréquences de début de période sont directement calculées à partir des données disponibles. Celles de fin de période sont estimées par extrapolation des séries antérieures reconstituées dans le tableau 1 de l'article de Jean-Paul Grémy [12, p. 399].

La propriété de variation en V garantit la diminution des indices lorsque les proportions se rapprochent par augmentation de la proportion inférieure ou par diminution de la proportion supérieure, et garantit leur augmentation lorsque la proportion supérieure augmente ou lorsque la proportion inférieure diminue. Elle établit une relation d'inégalité stricte entre les couples (x,y) de proportions distinctes lorsque les intervalles [x,y] sont liés par une relation d'inclusion stricte : $\forall(x,y)\in I^2$ tel que $x\neq y$ $\forall(u,v)\in I^2$ tel que $u\neq v$ [u,v] strictement inclus dans [x,y] $\Rightarrow M(u,v)<M(x,y)$. Cette propriété caractérise les indices qui, pour toute valeur donnée de l'une des proportions, varient de façon strictement monotone croissante en fonction de l'autre proportion si elle est supérieure à la première, et de façon strictement monotone décroissante si elle est inférieure. Si M est une fonction différentiable, la propriété peut s'exprimer : $\forall(x,y)\in I^2$ tel que $x\neq y$ $(x-y)\partial M/\partial_x(x,y)>0$ et $(y-x)\partial M/\partial_y(x,y)>0$. L'indice M a donc une variation en forme de V en fonction de chacune des proportions x et y.

La plupart des indicateurs utilisés dans les comparaisons d'inégalité entre proportions dispose des propriétés de positivité, d'identité, de symétrie et de variation en V. Ils constituent donc des indices d'inégalité. Tel est le cas notamment de l'*écart absolu E* défini par $E(x,y)=|x-y|$ [8, p.237 ; 15, pp.606 et 611 ; 20, p.395], du *rapport ordonné* R[8] défini par $R(x,y)=\mathrm{Max}\{x/y,y/x\}$ et de l'*indice logistique* L défini par $L(x,y)= |\mathrm{Log}(yx^*/xy^*)|$ [8, pp.246-248 ; 15, pp.622-624 ; 20, pp.397-401].

2.2. Propriété de cohérence

La propriété de cohérence[9] est introduite par Jean-Paul Grémy [12, p.410] et par Jean Prévot [15, p.605] pour garantir l'invariance des indices d'inégalité selon qu'ils portent sur les proportions x et y ou sur les proportions complémentaires $x^*=1-x$ et $y^*=1-y$. Elle s'exprime sous la forme : $\forall(x,y)\in I^2$ $M(x^*,y^*)=M(x,y)$.

Cette propriété concerne exclusivement les proportions selon lesquelles se répartissent des populations statistiques dans les deux modalités d'une variable dichotomique. Si on s'intéresse à l'inégalité entre les proportions d'accès à un certain niveau du système d'enseignement, la propriété de cohérence assure la compatibilité des résultats obtenus en comparant d'une part les indices d'inégalité entre les proportions d'admis et d'autre part les indices d'inégalité entre les proportions d'exclus.

Parmi les exemples courants d'indices d'inégalité cohérents, figurent l'écart absolu E et l'indice logistique L. Mais le rapport ordonné R n'en fait pas partie.

2.3. Propriété d'homogénéité

La propriété d'homogénéité est introduite dans la définition de certains indices d'inégalité pour en assurer l'invariance lorsque les variables sur lesquelles ils portent sont multipliées par un même scalaire positif [7, p.902 ; 16, p.620]. Ces indices sont donc des fonctions homogènes de degré 0 de ces variables. Dans le cas d'indices appliqués à deux proportions x et y, la propriété d'homogénéité s'exprime sous la forme : $\forall(x,y)\in I^2$ $\forall k>0$ tel que $(kx,ky)\in I^2$ $M(kx,ky)=M(x,y)$.

---

[8] Fréqemment utilisés [8, pp. 237-240], les rapports entre proportions ne constituent des indices d'inégalité qu'à condition d'être calculés en rapportant systématiquement la plus grande proportion à la plus petite. Leur minimum est égal à 1. Ils ont pour expression : $R(x,y)=\mathrm{Max}\{x,y\}/\mathrm{Min}\{x,y\}= \mathrm{Max}\{x/y,y/x\}$. Par la suite ces rapports sont qualifiés de rapports ordonnés, pour être distingués des rapports de la forme y/x, qui eux ne suivent pas les propriétés d'identité, de symétrie ou de variation en V.

[9] Jean-Paul Grémy ne qualifie pas cette propriété. Jean Prévôt l'appelle propriété de symétrie. Elle est qualifiée ici de propriété de cohérence pour éviter toute ambiguïté avec la propriété de symétrie introduite précédemment en d'autres termes.

Le rapport ordonné R constitue un exemple d'indice d'inégalité homogène, contrairement à l'écart absolu E ou à l'indice logistique L.

2.4. Incompatibilité entre cohérence et homogénéité des indices d'inégalité

Les indices d'inégalité ne peuvent pas être à la fois cohérents et homogènes.
Démonstration :

Soit M un indice d'inégalité défini sur I^2. On note repectivement (S) et (V) ses propriétés de symétrie et de variation en V. On suppose qu'il est muni des propriétés de cohérence (C) et d'homogénéité (H).

$\forall(x,y)\in I^2$ tel que $x<1$, $x\neq y$ et $x+y>1$

$x<1$ et $x+y>1 \Rightarrow 0<k=(1-x)/y<1 \Rightarrow (kx,ky)=((1-x)x/y,1-x)\in I^2$

(S) $\Rightarrow$ $M(x,y)=M(y,x)$

(H) $\Rightarrow$ $M(y,x)=M(ky,kx)=M[1-x,(1-x)x/y]$

(C) $\Rightarrow$ $M[1-x,(1-x)x/y]=M[x,1-(1-x)x/y]$

Soit $z=1-(1-x)x/y$

(S), (H) et (C) $\Rightarrow$ $M(x,z)=M(x,y)$

$z-x=(1-x)(y-x)/y$ et $z-y=(y-x)[1-(x+y)]/y$

$x<1$, $x\neq y$ et $x+y>1 \Rightarrow (z-x)(z-y)<0$ donc $z\in]x,y[$

$\forall(x,y)\in I^2$ tel que $x<1$, $x\neq y$ et $x+y>1$ $\exists z=1-(1-x)x/y$ strictement compris entre x et y tel que $M(x,z)=M(x,y)$. La propriété (V) n'est donc pas respectée. Un indice symétrique qui vérifie (V) ne peut donc pas respecter simultanément (C) et (H). A fortiori un indice d'inégalité ne peut pas être à la fois cohérent et homogène.

2.5. Indices d'inégalité relatifs

Les indices d'inégalité cohérents n'étant pas homogènes, les comparaisons de situations d'inégalité qu'ils permettent dépendent de l'ordre de grandeur des proportions qui sont confrontées. Il en résulte que l'appréciation de l'évolution des inégalités peut être fortement biaisée si les proportions connaissent elles-mêmes une variation importante. Ainsi, si on choisit l'écart absolu comme indice d'inégalité entre proportions, une augmentation forte des proportions risque de se répercuter sur l'accroissement de leur différence, donc d'entraîner un constat d'élévation des inégalités. Pour réduire ce biais, on est amené à définir des indices d'inégalité relatifs [13, pp.25-26 ; 14, p.634]. Par construction, de tels indices relativisent la prise en compte des proportions x et y qui décrivent une situation d'inégalité en les référant à une proportion r caractéristique de cette situation. Pour les proportions x et y d'admis parmi les enfants de cadres et les enfants d'ouvriers, on peut retenir comme proportion de référence la proportion m d'admis parmi l'ensemble des enfants de cadres et d'ouvriers ou la proportion t d'admis parmi les enfants de l'ensemble des n ($n\geq2$) catégories de la population active.

Comme les proportions de référence r varient en même temps que les couples (x,y) de proportions auxquelles elles se rapportent, les indices relatifs constituent des familles (M_r) paramétrées sur J (J=[0,1],]0,1], [0,1[ou]0,1[). Par définition une famille paramétrique (M_r) constitue une famille d'indices d'inégalité relatifs si quel que soit r dans J M_r est un indice d'inégalité, donc si M_r suit les propriétés de positivité, d'identité, de symétrie et de variation en V.

Parmi les familles d'indices d'inégalité relatifs d'usage courant, figure notamment la famille (ER_r) des *écarts relatifs* définis par $ER_r(x,y)=|x-y|/r$, où r est la proportion de l'ensemble de référence [8, p.245 ; 11, pp.36-37].

Pour les familles d'indices d'inégalité relatifs, la propriété de cohérence s'exprime en faisant intervenir la proportion de référence r et la proportion complémentaire r* : $\forall r \in J$ $\forall (x,y) \in I^2$ $M_{r^*}(x^*,y^*)=M_r(x,y)$.

Quant à la propriété d'homogénéité, elle se formule : $\forall r \in J$ $\forall (x,y) \in I^2$ $\forall k>0$ tel que $kr \in J$ et $(kx,ky) \in I^2$ $M_{kr}(kx,ky)=M_r(x,y)$. Elle est suivie notamment par les familles d'indices qui s'expriment comme fonctions des proportions relatives x/r et y/r. C'est en particulier le cas de la famille (ER_r) des écarts relatifs : $ER_r(x,y)=|x-y|/r=|x/r-y/r|$.

2.6. Indices d'inégalité pondérés

Jusqu'à présent, les indices d'inégalité entre proportions ont été définis sans référence à la pondération des catégories sur lesquels ils portent. Une telle démarche se fonde sur le présupposé que les unités à prendre en compte dans les comparaisons d'inégalité sont les catégories elles-mêmes. Mais si on considère que les unités à considérer dans les comparaisons ne sont pas les catégories, mais les individus de ces catégories, il convient d'utiliser des indices d'inégalité pondérés par les fréquences des catégories. En conséquence, pour des proportions fixées, les indices d'inégalité pondérés sont d'autant plus élevés que les catégories ont des fréquences équilibrées, et d'autant plus faibles que les catégories ont des fréquences disproportionnées. De ce point de vue, une société composée de privilégiés et de défavorisés est d'autant moins inégalitaire que le nombre de privilégiés (ou de défavorisés) est faible [10, pp.256-257].

Les indices d'inégalité entre proportions x et y pondérés par les fréquences a et a*=1-a des catégories correspondantes forment des familles (M_a) paramétrées sur]0,1[. Par définition ces familles sont des familles d'indices d'inégalité si quel que soit $a \in]0,1[$: d'une part M_a est muni des propriétés de positivité, d'identité et de variation en V; et d'autre part $\forall (x,y) \in I^2$ $M_{a^*}(y,x)=M_a(x,y)$.

En outre les familles d'indices d'inégalité pondérés sont cohérentes ou homogènes si quelle que soit la fréquence a les indices M_a respectent les propriétés de cohérence ou d'homogénéité.

2.7. Relation d'équivalence entre indices d'inégalité

Certains indices d'inégalité conduisent à des résultats analogues dans la comparaison de l'inégalité entre couples de proportions. On peut donc établir une relation d'équivalence entre indices produisant le même préordre entre couples de proportions.

On vérifie aisément que la transformation d'un indice d'inégalité M de minimum m_0 par une fonction f définie sur $[m_0,+\infty[$, strictement monotone croissante et telle que $f(m_0)$ soit positif ou nul, est un indice d'inégalité $N=f_oM$ équivalent à M. Il est doté des mêmes propriétés.

Il en résulte que les indices d'inégalité définis à un scalaire positif multiplicatif près sont équivalents entre eux. De même le rapport ordonné R est équivalent à l'*écart logarithmique* EL défini par $EL(x,y)=|Logx-Logy|$. En effet $EL(x,y)=|Logx-Logy|=|Log(x/y)|=$ $Log[Max\{x/y,y/x\}]=Log[R(x,y)]$. Comme la fonction logarithme est strictement monotone croissante et positive ou nulle sur $[1,+\infty[$ et comme 1 est le minimum de la fonction R, l'écart logarithmique EL et le rapport ordonné R sont équivalents.

3. INDICES D'INÉGALITÉ ASSOCIÉS A DES MESURES

L'indice d'inégalité cohérent d'usage le plus courant, l'écart absolu, offre l'avantage d'avoir une expression très simple. Mais il présente l'inconvénient d'accorder la même importance à des différences égales entre proportions quelle que soit la valeur de ces proportions. Il s'agit d'une hypothèse forte qui ne se conforme pas toujours aux présupposés théoriques que le sociologue ou l'économiste souhaite adopter pour interpréter les inégalités entre proportions. Ainsi, lorsque les variables étudiées sont supposées varier selon un modèle exponentiel, un écart de 10 points entre des proportions de 1% et de 11% s'avère beaucoup plus important qu'entre des proportions de 40% et de 50%. Dans de nombreux cas, il est donc plus pertinent d'attribuer à des différences égales entre proportions des valeurs distinctes qui tiennent compte de la valeur des proportions.

3.1. Définition d'indices d'inégalité à partir de mesures

Pour faire varier la valeur attribuée à des différences égales entre proportions, on peut transformer les différences infinitésimales Δx entre proportions en les pondérant par les valeurs $\mu(x)$ d'une fonction strictement positive définie sur I (I=[0,1],]0,1], [0,1[ou]0,1[). On obtient ainsi des différences infinitésimales $\mu(x)\Delta x$ qui dépendent de la valeur des proportions. L'intégration entre x et y de la fonction μ, supposée intégrable sur I, conduit à la définition d'un indice M de la forme : $M(x,y)=\left|\int_x^y \mu(u)du\right|$. C'est un indice d'inégalité. En effet il respecte les propriétés de positivité et de symétrie du fait de son expression en valeur absolue. Il suit la propriété d'identité en raison de la positivité stricte de la fonction μ. Enfin il satisfait à la propriété de variation en V : M étant différentiable par définition $\forall (x,y) \in I^2$ tel que $x \neq y$ $\partial M/\partial_x(x,y)=-\mu(x)<0$ si $x<y$ et $\partial M/\partial_x(x,y)=\mu(x)>0$ si $x>y$, donc $(x-y)\partial M/\partial_x(x,y)>0$; de la même façon $(y-x)\partial M/\partial_y(x,y)>0$.

Cette définition peut s'interpréter dans le cadre de la théorie de la mesure. En effet l'indice d'inégalité M construit à partir d'une fonction μ peut être associé à la mesure positive m de densité μ par rapport à la mesure de Lebesgue λ : $M(x,y)=m([x,y])=\int_{[x,y]}\mu d\lambda$. En particulier l'écart absolu correspond à la mesure de Lebesgue : $E(x,y)=\lambda([x,y])$. Par la suite on appelle densités les fonctions μ qui fondent la définition des indices d'inégalité associés à des mesures.

Les densités qui sont à la base de la définition de ces indices d'inégalité s'interprètent comme des modes de valorisation des différences élémentaires entre proportions. Leur détermination doit donc se conformer aux présupposés théoriques adoptés pour comparer et interpréter les inégalités. Ainsi l'écart absolu est obtenu avec une densité uniforme : $\left|\int_x^y du\right|=|y-x|=E(x,y)$. Il donne donc la même valeur aux différences élémentaires entre proportions, quelle que soit la valeur des proportions. L'écart logarithmique EL auquel est équivalent le rapport ordonné R est obtenu en choisissant la fonction inverse comme densité : $\left|\int_x^y u^{-1}du\right|=|Log(y/x)|=EL(x,y)=Log[R(x,y)]$. Il attribue donc aux différences élémentaires entre proportions des valeurs d'autant plus fortes que les proportions sont faibles. L'indice logistique L est obtenu à partir de la densité d'expression $\mu(x)=1/xx^*$: $\left|\int_x^y (uu^*)^{-1}du\right|=|Log(yx^*/xy^*)|=L(x,y)$. Compte tenu de la forme en U de la fonction de densité, la valorisation des différences élémentaires entre proportions est d'autant plus forte que les valeurs des proportions sont soit faibles, soit élevées.

Il en résulte que les résultats des comparaisons fondées sur les indices d'inégalité associés à des mesures dépendent du choix des densités. Ainsi, sur la base des données de scolarisation de l'exemple retenu, on constate que les inégalités d'accès à l'enseignement secondaire entre enfants de cadres et d'ouvriers évoluent de façon contradictoire selon les indices utilisés. En effet, à l'augmentation de l'écart absolu entre proportions d'admis (de 0,36 à 0,52), s'oppose la diminution du rapport ordonné (de 37 à 6,2), de l'écart logarithmique (de 3,61 à 1,82) et de l'indice logistique (de 4,06 à 2,69).

3.2. Condition de cohérence des indices d'inégalité de densité μ

Une condition nécessaire et suffisante pour qu'un indice d'inégalité défini à partir d'une densité μ soit cohérent est que la fonction μ soit symétrique par rapport à l'axe x=1/2. Cette condition s'exprime : $\forall x \in I\ \mu(x^*)=\mu(x)$.

Démonstration :

Soit M un indice d'inégalité de densité μ définie sur I=[0,1] ou]0,1[.

$$M(x,y)=\left|\int_x^y \mu(u)du\right|$$

$$M(x^*,y^*)=\left|\int_{x^*}^{y^*} \mu(u)du\right|=\left|-\int_x^y \mu(u^*)du\right|=\left|\int_x^y \mu(u^*)du\right|$$

La fonction μ étant positive strictement :

$$\forall (x,y)\in I^2\ \ M(x^*,y^*)=M(x,y) \iff \forall x\in I\ \ \mu(x^*)=\mu(x).$$

On vérifie ainsi à partir de la forme symétrique de leur fonction de densité que l'écart absolu et l'indice logistique sont des indices d'inégalité cohérents, contrairement aux écarts logarithmiques.

3.3. Condition d'homogénéité des indices d'inégalité de densité μ

Une condition nécessaire et suffisante pour qu'un indice d'inégalité défini à partir d'une densité μ soit homogène est que la fonction μ soit une fonction inverse, donc que l'indice soit proportionnel à l'écart logarithmique. A un scalaire multiplicatif près, l'écart logarithmique est donc le seul indice d'inégalité de densité μ à être homogène.

Démonstration :

Soit M un indice d'inégalité de densité μ définie sur I=]0,1].

$$M(x,y)=\left|\int_x^y \mu(u)du\right|$$

$$\forall k>0 \text{ tel que } (kx,ky)\in I^2 \quad M(kx,ky)=\left|\int_{kx}^{ky} \mu(u)du\right|=\left|\int_x^y k\mu(ku)du\right|$$

Supposons que M soit homogène

La fonction μ étant positive strictement :

$$\forall (x,y)\in I^2\ \forall k>0 \text{ tel que } (kx,ky)\in I^2\ \ M(kx,ky)=M(x,y) \Rightarrow k\mu(kx)=\mu(x)$$

$$k=1/x \Rightarrow \mu(x)=\mu(1)/x$$

$$M(x,y)=\left|\int_x^y \mu(u)du\right|=\left|\int_x^y [\mu(1)/u]du\right|=\mu(1).|Log(y/x)|=\mu(1).EL(x,y)$$

μ est une fonction inverse et M est proportionnel à l'écart logarithmique.

Supposons que M soit l'écart logarithmique : $M(x,y)=|Log(y/x)|$. Il admet la fonction inverse comme densité associée.

$$\forall (x,y)\in]0,1]^2\ \forall k>0 \text{ tel que } (kx,ky)\in]0,1]^2$$

$$M(kx,ky)=|Log(ky/kx)|=|Log(y/x)|=M(x,y)$$

M est homogène.

3.4. Indices d'inégalité de densité symétrique et convexe

Les densités associées aux indices d'inégalité usuels sont en général convexes. Si on munit simultanément les densités de la propriété de convexité conformément aux représentations courantes et de la propriété de symétrie par rapport à x=1/2 pour garantir la cohérence des indices, leur courbe représentative a une forme caractéristique de U symétrique par rapport à l'axe x=1/2.

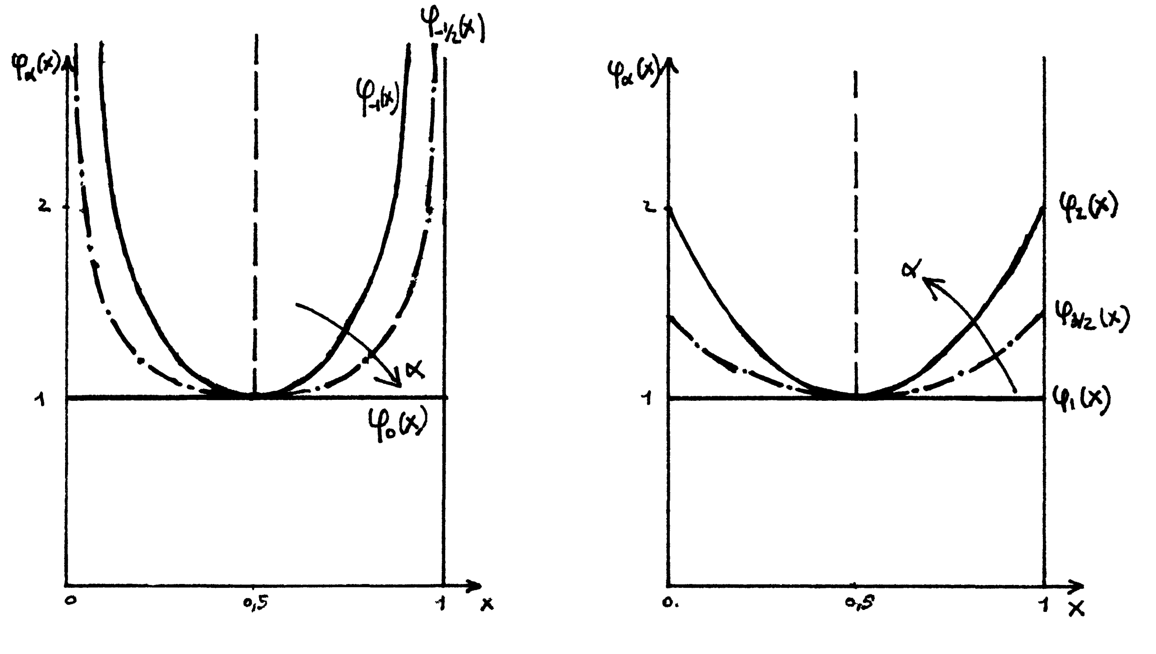
Parmi les densités symétriques et convexes figurent les fonctions de la forme $\mu(x)=f(x)+f(x^*)$, où f est une fonction positive et convexe. En effet, la fonction μ est symétrique par rapport à x=1/2 car elle s'exprime symétriquement en fonction de x et de x*. Par ailleurs, comme x* est l'image de x par la fonction affine $g(x^*=g(x)=1-x)$, $\mu=f+fog$. g étant affine et f convexe, fog est convexe. μ est donc convexe en tant que somme de deux fonctions convexes.

En particulier, le choix de fonctions f de puissance α ($\alpha\leq 0$ ou $\alpha\geq 1$) conduit à la définition d'une famille paramétrique de densités ϕ_α symétriques et convexes de la forme :

$$\phi_\alpha(x)=2^{\alpha-1}(x^\alpha+x*^\alpha)=[(2x)^\alpha+(2x^*)^\alpha]/2.$$

Avec des valeurs particulières du paramètre α on retrouve les densités des indices d'inégalité cohérents usuels. Ainsi ϕ_0 et ϕ_1 sont égales à la densité de l'écart absolu : $\phi_0(x)=\phi_1(x)=1$; et ϕ_{-1} est liée à la densité de l'indice logistique : $\phi_{-1}(x)=(1/xx^*)/4$.

Ces densités qui prennent toutes la valeur minimum 1 pour la proportion 1/2, ont une forme en U d'autant plus fermée que α est faible lorsqu'il est négatif, ou que α est élevé lorsqu'il est supérieur à 1, comme le montre les graphiques G1. Leur expression paramétrique permet donc de les choisir en les ajustant au mieux aux présupposés théoriques adoptés pour interpréter les inégalités.



Cas $\alpha\leq 0$ Cas $\alpha\geq 1$

Graphiques G1 : Variation des densités ϕ_α

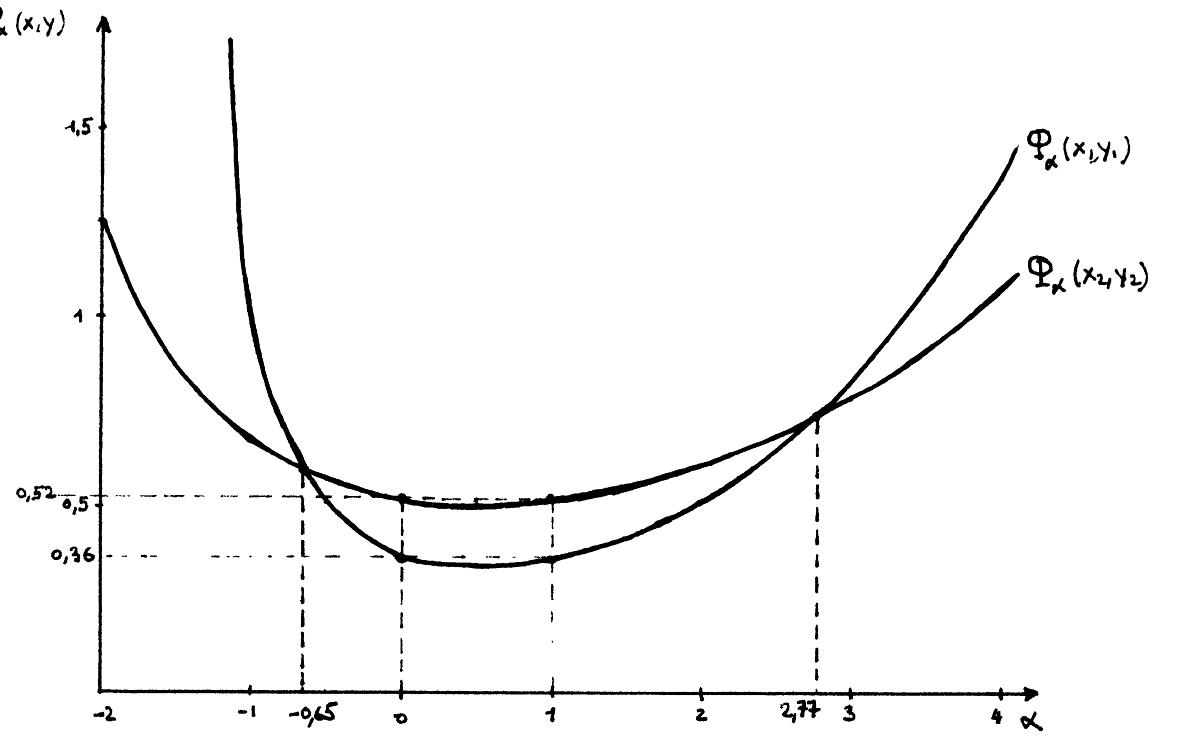
Ces densités permettent la construction d'une famille (Φ_α) d'indices d'inégalité de la forme :

$\Phi_\alpha(x,y)=(2^{\alpha-1}/|\alpha+1|)[|y^{\alpha+1}-x^{\alpha+1}|+|y*^{\alpha+1}-x*^{\alpha+1}|]$ si $\alpha\neq-1$;

$\Phi_{-1}(x,y)=(1/4)|Log(yx^*/xy^*)|$ si $\alpha=-1$.

Avec les valeurs 0 ou 1 du paramètre α, on obtient l'écart absolu ($\Phi_1=\Phi_0=E$). Avec la valeur -1, on obtient le quart de l'indice logistique ($\Phi_{-1}=L/4$).

Avec les données de l'exemple retenu, on constate pour les valeurs négatives ou nulles de α que les indices d'inégalité Φ_α augmentent si $\alpha>-0,65$ et diminuent si $\alpha<-0,65$, comme le montre le graphique G2. Et pour les valeurs de α supérieures ou égales à 1, les indices d'inégalité s'accroissent si $\alpha<2,77$ et décroissent si $\alpha>2,77$. On vérifie en particulier que l'écart absolu $E=\Phi_0=\Phi_1$ augmente et que l'indice logistique $L=4\Phi_{-1}$ diminue, comme il a été observé précédemment. Dans cet exemple, plus la forme en U des densités est fermée, plus l'indice d'inégalité associé tend à diminuer.



Graphique G2 : Évolution des indices Φ_α

4. FAMILLES D'INDICES D'INÉGALITÉ RELATIFS ASSOCIÉES A DES FAMILLES DE MESURES PARAMÉTRIQUES

4.1. Définition de familles d'indices d'inégalité relatifs à partir de familles de mesures paramétriques

Jusqu'à présent, les densités utilisées pour construire des indices d'inégalité entre les proportions de deux catégories ont été définies comme des fonctions ne variant qu'avec les proportions. Mais dans certains cas il est souhaitable que la valorisation donnée aux différences élémentaires entre proportions dépende non seulement de la valeur des proportions, mais aussi de celle d'une proportion de référence commune, par exemple la proportion moyenne observée dans l'ensemble des catégories. Ainsi on peut donner une signification plus importante à un écart de 1 point entre deux proportions de 10 et de 11% lorsque la proportion moyenne est de 5% que lorsqu'elle est de 20%, comme le suggère l'écart relatif qui s'élève à 20% dans le premier cas et seulement à 5% dans le second.

Pour relativiser les indices d'inégalité par rapport à des proportions de référence r, on est conduit à adopter des densités qui en dépendent. Ces densités peuvent être définies sur I (I=[0,1],]0,1], [0,1[ou]0,1[) sous la forme paramétrique μr, r variant sur J (J=[0,1],]0,1], [0,1[ou]0,1[). Les familles de densités μ_r conduisent à la définition de familles d'indices d'inégalité relatifs M_r de la forme : $M_r(x,y)=\left|\int_x^y \mu_r(u)du\right|$. Ces familles (M_r) sont associées aux familles de mesures positives de densités μ_r par rapport à la mesure de Lebesgue. En particulier, avec des densités 1/r, on obtient la famille des écarts relatifs : $\left|\int_x^y r^{-1}du\right|=|y-x|/r=ER_r(x,y)$.

De même que le choix de l'indicateur, le choix de la proportion de référence joue un rôle important dans la définition des indices d'inégalité relatifs et peut conduire à des conclusions divergentes. Ainsi, avec les données de l'exemple retenu, on observe que l'écart relatif entre proportions d'admis diminue nettement (de 3 à 2,26) si on le calcule en référence à la proportion moyenne d'admis de l'ensemble des catégories sociales, alors qu'il reste stable à la valeur 2 si on le calcule en référence à la proportion moyenne d'admis des deux seules catégories ouvriers et cadres. Différentes entre elles, ces évolutions d'écarts relatifs divergent aussi par rapport à l'augmentation observée avec l'écart absolu. Elles diffèrent également de celles qu'on observe avec les écarts relatifs entre proportions d'exclus qui croissent respectivement de 0,41 à 0,68 et de 0,44 à 0,70 selon qu'ils sont rapportés aux proportions moyennes globales ou partielles.

4.2. Condition de cohérence des indices d'inégalité relatifs de densités μ_r

Une condition nécessaire et suffisante pour qu'une famille d'indices d'inégalité relatifs définie à partir d'une famille de densités μr soit cohérente est qu'il y ait égalité entre les densités des proportions complémentaires référées à r* et les densités des proportions référées à r : $\forall r \in J\ \forall x \in I\ \mu_r{}^*(x^*)=\mu_r(x)$. Les densités respectant cette condition sont dites bisymétriques (double symétrie par rapport à x=1/2 et par rapport à r=1/2).
Démonstration :

Soit (M_r) la famille d'indices d'inégalité relatifs associée à la famille des densités μ_r définies sur I=[0,1] ou]0,1[et paramétrées sur J=[0,1] ou]0,1[

$M_r(x,y)=\left|\int_x^y \mu_r(u)du\right|$

$M_{r^*}(x^*,y^*)=\left|\int_{x^*}^{y^*} \mu_{r^*}(u)du\right|=\left|-\int_x^y \mu_{r^*}(u^*)du\right|=\left|\int_x^y \mu_{r^*}(u^*)du\right|$

Les fonctions μ_r étant positives strictement :

$\forall r \in J\ \forall (x,y) \in I^2\ \ M_r{}^*(x^*,y^*)=M_r(x,y) \iff \forall r \in J\ \forall x \in I\ \ \mu_{r^*}(x^*)=\mu_r(x)$.

Il en résulte que la famille des écarts relatifs n'est pas cohérente puisqu'elle est définie avec des densités non bisymétriques, de la forme 1/r.

4.3. Condition d'homogénéité des indices d'inégalité relatifs de densités μ_r

Une condition nécessaire et suffisante pour qu'une famille d'indices d'inégalité relatifs définie à partir d'une famille de densités μ_r soit homogène est que les fonctions μ_r vérifient la relation : $\forall r \in J\ \forall x \in I\ \forall k>0$ tel que $kx \in I$ et $kr \in J$ $\mu_{kr}(kx)=\mu_r(x)/k$. De plus, si les densités μ_r sont les restrictions sur I de fonctions π_r vérifiant cette relation sur $]0,+\infty[$, les indices d'inégalité $M_r(x,y)$ s'expriment en fonction des proportions relatives x/r et y/r.

Démonstration :

Soit (M_r) la famille d'indices d'inégalité relatifs associée à la famille des densités μ_r définies sur $I=]0,1]$ et paramétrées sur $J=]0,1]$.

$M_r(x,y)=|\int_x^y \mu_r(u)du|$

$\forall k>0$ tel que $kr\in J$ et $(kx,ky)\in I^2$ $M_{kr}(kx,ky)=|\int_{kx}^{ky} \mu_{kr}(u)du|=|\int_x^y k\mu_{kr}(ku)du|$

Supposons que les indices M_r soient homogènes.

Les fonctions μ_r étant positives strictement : $\forall r\in J$ $\forall(x,y)\in I^2$ $\forall k>0$ tel que $kr\in J$ et $(kx,ky)\in I^2$ $M_{kr}(kx,ky)=M_r(x,y) \Leftrightarrow k\mu_{kr}(kx)=\mu_r(x)$

Supposons que les densités μ_r soient les restrictions sur I de fonctions π_r définies sur $]0,+\infty[$ et vérifiant la relation :

$\forall r\in J$ $\forall x>0$ $\forall k>0$ tel que $kr\in J$ $k\pi_{kr}(kx)=\pi_r(x)$

$\forall r\in J$ $\forall x\in I$ $k=1/r \rightarrow \mu_r(x)=\pi_r(x)=\pi_1(x/r)/r$

$M_r(x,y)=|\int_x^y \mu_r(u)du|=|\int_x^y [\pi_1(u/r)/r]du|=|\int_{x/r}^{y/r} \pi_1(u)du|$

$M_r(x,y)$ s'exprime donc en fonction des proportions relatives x/r et y/r.

4.4. Indices d'inégalité relatifs de densité bisymétrique et convexe

Les propriétés de forme des densités présentées au paragraphe 3.4, s'étendent au cas des indices d'inégalité relatifs. Ainsi à la propriété de bisymétrie des densités qui assure la cohérence des indices, il est conforme aux représentations courantes d'ajouter celle de convexité.

Parmi les densités bisymétriques et convexes associées à des indices d'inégalité relatifs, figure la famille des densités $\psi_{r\alpha}$ qui généralisent les densités ϕ_α définies au paragraphe 3.4. Elles s'expriment sous la forme :

$$\psi_{r\alpha}(x)=[(x/r)^\alpha+(x^*/r^*)^\alpha]/2 \quad (\alpha\leq 0 \text{ ou } \alpha\geq 1).$$

Les densités ϕ_α en constituent le sous-ensemble correspondant au cas r=1/2 : $\phi_\alpha=\psi_{1/2,\alpha}$.

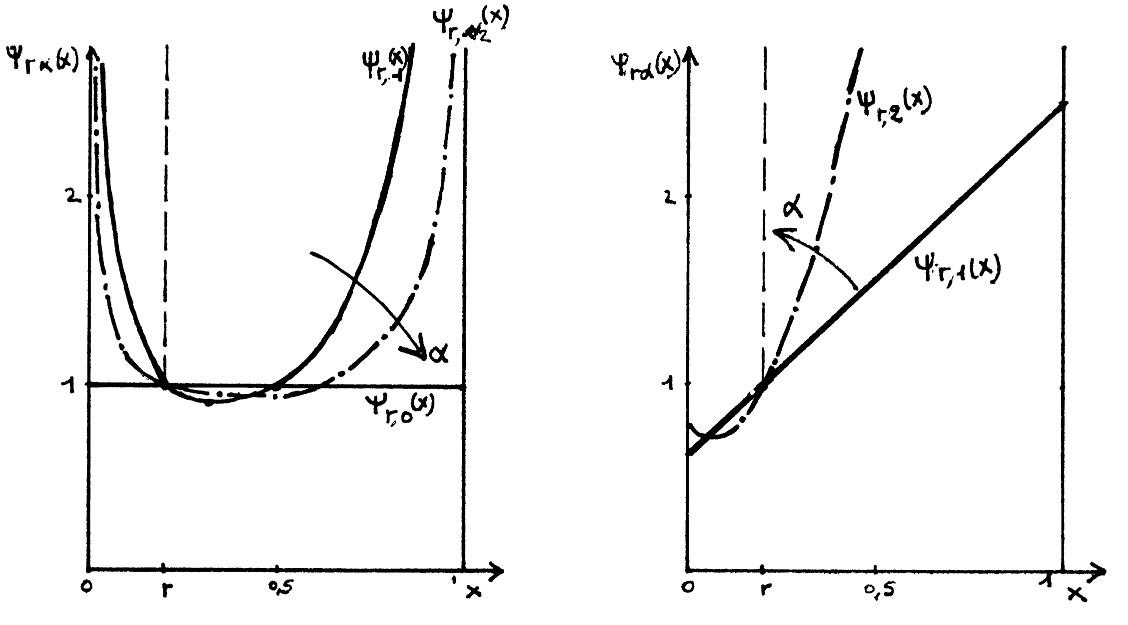
Les densités $\psi_{r\alpha}$ (voir graphique G3 page suivante) ont une forme en U si $\alpha\leq 0$ ou en J si $\alpha\geq 1$ comme le montrent les graphiques G3. Elles atteignent leur minimum avec la proportion $p_\alpha(r)=1/[1+(r^*/r)^{\alpha/\alpha-1}]$, en général distincte de r. Lorsque $\alpha<0$, $p_\alpha(r)$ est comprise entre r et 1/2. Elle tend vers 1/2 si α tend vers 0 et vers r si α tend vers $-\infty$. Mais lorsque $\alpha>1$, le minimum est atteint pour des proportions extérieures à l'intervalle borné par r et 1/2. On aboutit donc à une représentation non conforme aux présupposés théoriques les plus courants.

Ces densités permettent de construire des familles d'indices d'inégalité relatifs $\Psi_{r\alpha}$:

$\Psi_{r\alpha}(x,y)=[r|(y/r)^{\alpha+1}-(x/r)^{\alpha+1}|+r^*|(y^*/r^*)^{\alpha+1}-(x^*/r^*)^{\alpha+1}|]/2|\alpha+1|$ si $\alpha\neq-1$;

$\Psi_{r,-1}(x,y)=[r|Log(y/x)|+r^*|Log(x^*/y^*)|]/2$ si $\alpha=-1$.

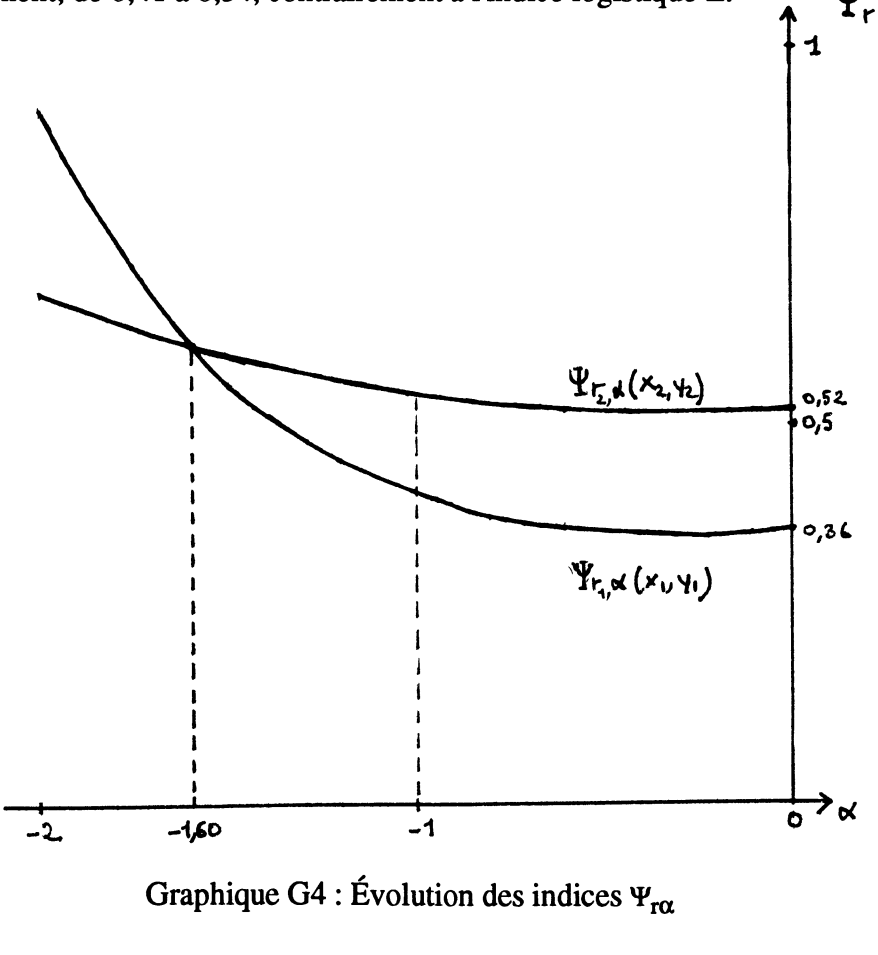
Avec la valeur 0 du paramètre α on retrouve l'écart absolu : $E=\Psi_{r0}$. Avec la valeur -1 on définit l'*indice logistique relatif* LR_r: $LR_r=\Psi_{r,-1}$. Par ailleurs les indices Φ_α sont obtenus avec r=1/2 : $\Phi_\alpha=\Psi_{1/2,\alpha}$.



Cas α≤0 Cas α≥1

Graphiques G3 : Variation des densités $\psi_{r\alpha}$

Avec les données de l'exemple retenu et en choisissant comme proportion de référence la proportion moyenne d'admis dans l'ensemble des catégories sociales, on constate que les indices d'inégalité relatifs $\Psi_{r\alpha}$ définis avec des valeurs négatives ou nulles de α augmentent si $\alpha>-1{,}60$ et diminuent si $\alpha<-1{,}60$, comme le montre le graphique G4. On vérifie en particulier que l'écart absolu $E=\Psi_{r0}$ augmente de 0,36 à 0,52 et que l'indice logistique relatif $LR_r=\Psi_{r,-1}$ s'accroît également, de 0,41 à 0,54, contrairement à l'indice logistique L.



Graphique G4 : Évolution des indices $\Psi_{r\alpha}$

5. INDICES D'INÉGALITÉ DÉFINIS A PARTIR DE DISTANCES

5.1. Définition d'indices d'inégalité comme distances fonctionnelles entre proportions

Les distances entre proportions permettent de définir des indices d'inégalité. C'est en particulier le cas de la distance d définie sur I^2 par $d(x,y)=|x-y|$, puisqu'elle est égale à l'écart absolu. Plus généralement, il en est de même pour les distances fonctionnelles d_f définies sur I^2 à partir de la distance d et de fonctions f strictement monotones croissantes sur I, sous la forme : $d_f(x,y)=d[f(x),f(y)]=|f(x)-f(y)|$. En effet elles respectent d'une part les propriétés de positivité, d'identité et de symétrie en tant que distances entre proportions. Et elles suivent d'autre part la propriété de variation en V : quel que soit z strictement compris entre x et y distincts, $d_f(x,y)$ est la somme des distances strictement positives $d_f(x,z)$ et $d_f(z,y)$, donc elle est supérieure à chacune d'elles.

L'utilisation d'une distance fonctionnelle d_f plutôt que la distance d comme indice d'inégalité revient à effectuer une transformation de l'échelle des proportions pour attribuer des significations différentes à des différences égales entre proportions, selon une démarche proposée par Jean-Paul Grémy [12, pp.405-410]. Au lieu de porter directement sur la différence entre les proportions, l'indice d'inégalité porte alors sur la différence entre leurs images par la fonction f de transformation d'échelle.

Lorsque la fonction f de transformation d'échelle est dérivable, la distance fonctionnelle d_f correspond à la mesure m_f de densité f' (dérivée de f) par rapport à la mesure de Lebesgue : $d_f(x,y)=|f(x)-f(y)|=|\int_x^y f'(u)du|=\int_{[x,y]} f'd\lambda=m_f([x,y])$. La cohérence et l'homogénéité des indices d'inégalité définis comme distances fonctionnelles d_f dépendent alors des propriétés de la densité f' (cf. 3.2 et 3.3).

Réciproquement les indices d'inégalité associés à des mesures de densité μ par rapport à la mesure de Lebesgue peuvent être considérés comme des distances fonctionnelles d_f de fonction f primitive de μ.

5.2. Définition d'indices d'inégalité à partir de distances fonctionnelles entre distributions de proportions dichotomiques

Les indices d'inégalité définis comme distances fonctionnelles entre proportions ne sont pas nécessairement cohérents. Mais ils le deviennent systématiquement s'ils sont définis à partir de certaines distances fonctionnelles entre les distributions de proportions dichotomiques $X=(x,x^*)$ et $Y=(y,y^*)$. C'est notamment le cas avec les distances fonctionnelles D_{fp}, de fonction f strictement monotone croissante sur I et d'ordre p, entre les distributions X et Y. Elles se définissent sur I^4 ($I=[0,1]$ ou $]0,1[$) à partir des distances de Minkowski d'ordre p entre les couples $(f(x),f(x^*))$ et $(f(y),f(y^*))$ sous la forme : $D_{fp}(X,Y)=[|f(x)-f(y)|^p+|f(x^*)-f(y^*)|^p]^{1/p}$ ($p\geq1$). Elles s'expriment en fonction des distances fonctionnelles d_f entre proportions : $D_{fp}(X,Y)=[d_f^p(x,y)+d_f^p(x^*,y^*)]^{1/p}$. Elles permettent de définir des indices d'inégalité cohérents de la forme : $M_{fp}(x,y)=D_{fp}(X,Y)$.

Démonstration :

$M_{fp}(x,y)=D_{fp}(X,Y)=[d_f^p(x,y)+d_f^p(x^*,y^*)]^{1/p}$

Positivité : $M_{fp}(x,y)=D_{fp}(X,Y)\geq0$

Identité : $M_{fp}(x,y)=0 \Leftrightarrow D_{fp}(X,Y)=0 \Leftrightarrow X=Y \Leftrightarrow x=y$

Symétrie : $M_{fp}(y,x)=D_{fp}(Y,X)=D_{fp}(X,Y)=M_{fp}(x,y)$

Variation en V : d_f vérifiant la propriété de variation en V

$\forall z \in]x,y[\ z^* \in]x^*,y^*[$ et : $d_f(x,z)<d_f(x,y)$ et $d_f(x^*,z^*)<d_f(x^*,y^*)$

$[d_f^p(x,z)+d_f^p(x^*,z^*)]^{1/p} < [d_f^p(x,y)+d_f^p(x^*,y^*)]^{1/p}$

$M_{fp}(x,z)<M_{fp}(x,y)$

On démontre de même que $M_{fp}(z,y)<M_{fp}(x,y)$

Cohérence : $M_{fp}(x^*,y^*)=[d_f^p(x^*,y^*)+d_f^p(x,y)]^{1/p}=$
$[d_f^p(x,y)+d_f^p(x^*,y^*)]^{1/p}=M_{fp}(x,y)$

Parmi les distances fonctionnelles d'ordre p entre distributions de proportions dichotomiques figurent notamment les distances $D_{\alpha p}$ construites avec des fonctions de puissance α ($\alpha \neq 0$) et la distance D_{0p} construite avec une fonction logarithme. Les indices d'inégalité associés s'expriment respectivement :

$M_{\alpha p}(x,y)=[|x^\alpha-y^\alpha|^p+|x^{*\alpha}-y^{*\alpha}|^p]^{1/p}$ ($\alpha \neq 0$)

$M_{0p}(x,y)=[|Log(x/y)|^p+|Log(x^*/y^*)|^p]^{1/p}$.

Les indices définis à partir des distances D_{1p} de fonction linéaire ($\alpha=1$) et d'ordre p sont tous proportionnels à l'écart absolu : $D_{1p}(X,Y)=[|x-y|^p+|x^*-y^*|^p]^{1/p}=2^{1/p}|x-y|= 2^{1/p}E(x,y)$. Parmi les distances rectilinéaires $D_{\alpha 1}$ (d'ordre p=1), les distances linéaire D_{11} ($\alpha=1$) et quadratique D_{21} ($\alpha=2$) correspondent au double de l'écart absolu. La distance logarithmique D_{01} ($\alpha=0$) définit l'indice logistique. Plus généralement les distances rectilinéaires $D_{\alpha 1}$ sont liées aux indices Φ_α : $D_{\alpha 1}(X,Y)=(|\alpha|/2^{\alpha-2})\Phi_\alpha\text{-}1(x,y)$ si $\alpha \neq 0$; et $D_{01}(X,Y)=4\Phi\text{-}1(x,y)$. Parmi les distances euclidiennes $D_{\alpha 2}$, la distance linéaire D_{12} ($\alpha=1$) génère l'écart absolu : $D_{12}(X,Y)=\sqrt{2}E(x,y)$; et la distance $D_{1/2,2}$ de fonction racine carrée ($\alpha=1/2$) correspond à l'*indice d'Hellinger* H [2 , p.37] : $D_{1/2,2}=\sqrt{H}$.

Les indices d'inégalité entre proportions définis à partir de distances fonctionnelles rectilinéaires D_{f1} entre distributions de proportions dichotomiques peuvent aussi s'exprimer comme des distances fonctionnelles d_g entre proportions, la fonction g se déduisant de f par la relation $g(x)=f(x)-f(x^*)$. En effet : $M_{f1}(x,y)=D_{f1}(X,Y)=|f(x)-f(y)| +|f(x^*)-f(y^*)|=|[f(x)-f(x^*)]-[f(y)-f(y^*)]|=d_g(x,y)$. En outre, si les fonctions f sont dérivables, il en résulte que ces indices peuvent aussi être construits à partir de densités μ de la forme $\mu(x)=g'(x)=f'(x)+f'(x^*)$. Par contre on ne peut pas établir une telle correspondance avec les indices construits à partir de distances fonctionnelles euclidiennes ou plus généralement d'ordre $p \neq 1$ définies avec des fonctions f non affines. Ces distances génèrent donc de nouvelles classes d'indices, distinctes de celles présentées dans les chapitres précédents.

Par ailleurs, à partir de familles de fonctions f_r strictement monotones croissantes sur I et paramétrées sur J, on peut définir des familles d'indices d'inégalité relatifs sur la base des familles de distances fonctionnelles de fonctions f_r entre proportions ou des familles de distances fonctionnelles de fonctions f_r et d'ordre p entre distributions de proportions dichotomiques.

6. INDICES D'INÉGALITÉ PONDÉRÉS DÉFINIS PARTIR DE CARACTÉRISTIQUES STATISTIQUES DE DISPERSION

Les différents indices d'inégalité qui ont été définis précédemment à partir de mesures ou de distances ne prennent pas en compte la pondération des catégories sociales qu'ils comparent. Pour les prendre en considération, il faut faire appel à d'autres types d'indicateurs. Parmi eux figurent notamment les caractéristiques statistiques de dispersion comme les écarts moyens étudiés dans ce chapitre, ainsi que les indices d'Atkinson et les fonctions et indices de concentration de Gini étudiés dans les chapitres suivants.

6.1. Définition d'indices d'inégalité pondérés comme écarts moyens des proportions

Appliquées à des couples de proportions pondérées par les fréquences des catégories sociales correspondantes, les caractéristiques statistiques de dispersion permettent de définir des indices d'inégalité pondérés. Ainsi les écarts moyens d'ordre p (p étant un entier positif) par rapport à la moyenne, des proportions x et y pondérées par les fréquences a et a*, constituent des indices pondérés E_{pa} de paramètre a, de la forme : $E_{pa}(x,y)=(aa^{*p}+a^{*}a^{p})^{1/p}|x-y|$. Ces indices sont proportionnels à l'écart absolu E (et à la distance d) entre proportions, avec comme coefficient de proportionnalité le facteur de pondération $k_p(a)=(aa^{*p}+a^{*}a^{p})^{1/p}$. En particulier l'écart-type E_{2a} et l'écart absolu moyen E_{1a} s'expriment : $E_{2a}=\sqrt{(aa^*)}E$ et $E_{1a}=2aa^*E$. Les coefficients $k_p(a)$ étant des fonctions symétriques en a et a*, les indices E_{pa} constituent des familles d'indices d'inégalité cohérents qui généralisent l'écart absolu en lui affectant un facteur de pondération.
Démonstration :

Soit un couple de proportions x et y pondérées par les fréquences a et a*

Soit m la moyenne pondérée de x et y : $m=ax+a^*y$

$x-m=x-(ax+a^*y)=a^*(x-y)$

$y-m=y-(ax+a^*y)=a(y-x)$

$$E_{pa}(x,y) =[a|x-m|^p+a^*|y-m|^p]^{1/p}=[a(a^*|x-y|)^p+a^*(a|y-x|)^p]^{1/p}$$
$$=(aa^{*p}+a^*a^p)^{1/p}|x-y|=(aa^{*p}+a^*a^p)^{1/p}E(x,y)$$

En particulier : $E_{1a}=2aa^*E$

$E_{2a}=\sqrt{(aa^*)}E.$

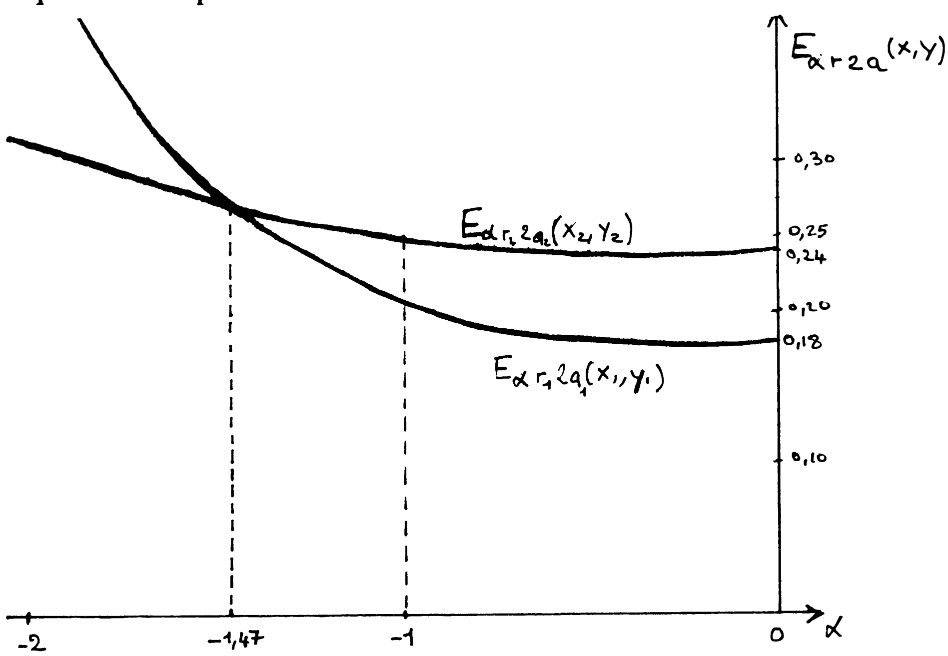
Avec les données de l'exemple retenu, on observe une augmentation des écarts moyens d'ordre p, de même qu'avec l'écart absolu. Mais les accroissements de l'écart-type (de 0,18 à 0,24) et plus encore de l'écart absolu moyen (de 0,18 à 0,22) présentent une moindre ampleur que celui de l'écart absolu, en raison de la diminution des facteurs de pondération qui résulte de la disproportion croissante entre les effectifs des enfants de cadres et d'ouvriers.

6.2. Définition d'indices d'inégalité pondérés comme écarts fonctionnels moyens des proportions

La relation qui vient d'être établie entre écarts moyens d'ordre p et écart absolu entre proportions vaut pour tout couple de valeurs d'une variable pondérée par les fréquences a et a*. Elle se vérifie donc pour les couples (f(x),f(y)) d'images des proportions par des fonctions f strictement monotones croissantes. En conséquence les écarts fonctionnels moyens d'ordre p des proportions E_{fpa} définis comme écarts moyens d'ordre p par rapport à la moyenne, des couples (f(x),f(y)), sont liés aux distances fonctionnelles df entre proportions, et si f est dérivable aux indices de densité f'. Ils les généralisent en y introduisant les facteurs de pondération $k_p(a)=(aa^{*p}+a^*a^p)^{1/p}$: $E_{fpa}(x,y)= E_{pa}(f(x),f(y))=k_p(a)|f(x)-f(y)|=k_p(a)df(x,y)$. Les écarts fonctionnels moyens d'ordre p constituent donc des familles d'indices d'inégalité pondérés ayant les mêmes propriétés de cohérence ou d'homogénéité que les distances fonctionnelles qui leur correspondent.

De même, les écarts fonctionnels moyens d'ordre p E_{f_rpa}, de fonctions f_r paramétrées par des proportions de référence r, forment des familles d'indices d'inégalité relatifs proportionnels aux distances fonctionnelles d_{f_r}, et si les fonctions f_r sont dérivables aux indices de densités f'_r. En particulier on peut associer les familles d'écarts fonctionnels moyens ($E_{\alpha rpa}$) aux familles d'indices d'inégalité relatifs ($\Psi_{r\alpha}$) définies au chapitre 4 : $E_{\alpha rpa}=k_p(a)\Psi_{r\alpha}$. Ces écarts forment des familles pondérées d'indices d'inégalité cohérents. Les familles ($\Psi_{r\alpha}$) et (Φ_α) en constituent des sous-ensembles. Les indices $\Psi_{r\alpha}$ s'en déduisent en neutralisant le paramètre de pondération a avec la valeur 1/2; et les indices Φ_α en neutralisant en outre le paramètre r caractérisant les proportions de référence avec la valeur 1/2.

Avec les données de l'exemple retenu et en choisissant comme proportions de référence r les proportions moyennes d'admis dans l'ensemble des catégories sociales, on constate que les écarts-type $E_{\alpha r2a}$ définis avec des valeurs négatives ou nulles de α, augmentent si α>-1,47 et diminuent si α<-1,47, comme le montre le graphique G5. On vérifie en particulier que l'écart-type E_{0r2a}, qui est égal à l'écart-type des proportions E_{2a}, augmente de 0,18 à 0,24; et que l'écart-type E_{-1r2a} passe de 0,21 à 0,24, donc augmente comme l'indice logistique relatif $LR_r=\Psi_{r,-1}$ auquel il correspond.



Graphique G5 : Evolution des écarts-type $E_{\alpha r2a}$

7. INDICES D'INÉGALITÉ PONDÉRÉS DÉFINIS A PARTIR DES INDICES D'ATKINSON

Les *indices d'Atkinson* [1], ou les indices additivement séparables et décomposables qui leur sont équivalents [7, p.917; 16, p.622], constituent une autre famille d'indices pouvant prendre en compte la pondération des catégories qui sont comparées. Dans le cas d'un couple de proportions x et y pondérées par les fréquences a et a* et de moyenne m=ax+a*y, et sous une forme additivement séparable, ils ont pour expression :

$$S_\alpha(x,y)=[a(x/m)^\alpha+a^*(y/m)^\alpha-1]/\alpha(\alpha-1) \qquad (\alpha\neq 0 \text{ et } \alpha\neq 1)$$

$$S_0(x,y)=\lim_{\alpha\to 0} S_\alpha(x,y)=-a\mathrm{Log}(x/m)-a^*\mathrm{Log}(y/m)$$

$$S_1(x,y)=\lim_{\alpha\to 1} S_\alpha(x,y)=a(x/m)\mathrm{Log}(x/m)+a^*(y/m)\mathrm{Log}(y/m)$$

Un certain nombre d'indices usuels se rattache à cette famille. Figurent parmi eux l'*indice entropique de Theil* $T=S_1$ [17, pp.126-127], la *déviation logarithmique moyenne* $DL=S_0$, le *coefficient de variation* CV ($CV^2=2S_2$) ou l'*indice d'Hellinger* H ($H=mS_{1/2}/2$)[10].

---

[10] Pour des proportions x et y, pondérées par les fréquences a et a* et de moyenne m=ax+a*y, l'indice d'Hellinger H et l'indice $S_{1/2}$ ont pour expression :
$H(x,y)=a(\sqrt{x}-\sqrt{m})^2+a^*(\sqrt{y}-\sqrt{m})^2=2m[1-(a\sqrt{(x/m)}+a^*\sqrt{(y/m)})]$
$S_{1/2}(x,y)=4[1-(a\sqrt{(x/m)}+a^*\sqrt{(y/m)})]$. D'où la relation : $H=mS_{1/2}/2$.

En tant que fonctions des proportions x et y paramétrées par la fréquence a, les indices de la famille (S_α) suivent par définition les propriétés de positivité, d'identité, de symétrie et d'homogénéité. Comme en outre ils vérifient la propriété de variation en V, ils constituent des familles pondérées d'indices d'inégalité homogènes.
Démonstration :

$$S_\alpha(x,y) = [a(x/m)^\alpha + a^*(y/m)^\alpha - 1]/\alpha(\alpha-1) \quad (\alpha \neq 0 \ \text{ et } \ \alpha \neq 1)$$
$$= [(ax^\alpha + a^*y^\alpha)/(ax+a^*y)^\alpha - 1]/\alpha(\alpha-1)$$
$$\partial S_\alpha/\partial x(x,y) = aa^*y(x^{\alpha-1} - y^{\alpha-1})/(\alpha-1)(ax+a^*y)^{\alpha+1}$$
$$S_0(x,y) = -a\mathrm{Log}(x/m) - a^*\mathrm{Log}(y/m)$$
$$= -a\mathrm{Log}x - a^*\mathrm{Log}y + \mathrm{Log}(ax+a^*y)$$
$$\partial S_0/\partial x(x,y) = -aa^*y(x^{-1} - y^{-1})/(ax+a^*y)$$
$$S_1(x,y) = a(x/m)\mathrm{Log}(x/m) + a^*(y/m)\mathrm{Log}(y/m)$$
$$= (ax\mathrm{Log}x + a^*y\mathrm{Log}y)/(ax+a^*y) - \mathrm{Log}(ax+a^*y)$$
$$\partial S_1/\partial x(x,y) = aa^*y(\mathrm{Log}x - \mathrm{Log}y)/(ax+a^*y)^2$$

$\forall\alpha\ \forall(x,y)\in]0,1[^2$ tel que $x \neq y$ $(x-y)\partial S_\alpha/\partial x(x,y) > 0$

De même : $(y-x)\partial S_\alpha/\partial y(x,y) > 0$

Les indices S_α vérifient la propriété de variation en V.

Par contre les indices S_α ne respectent pas la propriété de cohérence. En conséquence, pour comparer de façon cohérente les inégalités entre les couples de proportions (x_1,y_1) et (x_2,y_2), il convient de faire une double comparaison entre les indices $S_\alpha(x_1,y_1)$ et $S_\alpha(x_2,y_2)$ d'une part et entre les indices portant sur les proportions complémentaires $S_\alpha(x_1^*,y_1^*)$ et $S_\alpha(x_2^*,y_2^*)$ d'autre part. S'ils augmentent (respectivement diminuent) simultanément, les inégalités augmentent (respectivement diminuent). S'ils varient en sens inverse, on ne peut pas conclure. Une autre façon de procéder consiste à définir de nouveaux indices d'inégalité $T_{r\alpha}$ comme moyennes pondérées des indices d'inégalité S_α portant sur les proportions et sur les proportions complémentaires : $T_{r\alpha}(x,y) = rS_\alpha(x,y) + r^*S_\alpha(x^*,y^*)$. De tels indices sont alors cohérents par définition, mais ne sont pas homogènes. Lorsque les coefficients de pondération r et r* sont les proportions moyennes m=ax+a*y et m*= ax*+a*y* des deux catégories étudiées, ces indices constituent des *f-divergences* de Csiszar, indices de dépendance proposés par Jean-Pierre Florens comme indices d'inégalité cohérents [10, pp.258-262].

Avec les données de l'exemple retenu, les indices d'inégalité S_α appliqués aux proportions d'admis diminuent si $\alpha<2{,}6$ et augmentent si $\alpha>2{,}6$, comme le montre le tableau T3. Par contre, s'ils sont appliqués aux proportions d'exclus, tous les indices augmentent. Ainsi pour les valeurs du paramètre α supérieures à 2,6 les indices d'inégalité S_α relatifs aux admis et aux exclus présentent des résultats cohérents entre eux. Dans ce cas on peut donc conclure à un accroissement des inégalités.

α	0	1/2	1	2	2,6	3
$S_\alpha(x_1,y_1)$	1,19	0,83	0,59	0,47	0,48	0,50
$S_\alpha(x_2,y_2)$	0,40	0,39	0,39	0,43	0,48	0,54
$S_\alpha(x_1^*,y_1^*)$	0,02	0,01	0,03	0,03	0,03	0,03
$S_\alpha(x_2^*,y_2^*)$	0,07	0,07	0,06	0,05	0,05	0,05

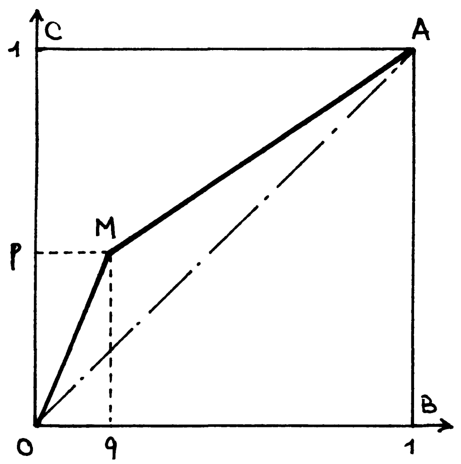
Tableau T3
Evolution des indices d'inégalité S_α

8. COMPARAISON DES INÉGALITES A PARTIR DES FONCTIONS ET DES INDICES DE CONCENTRATION DE GINI

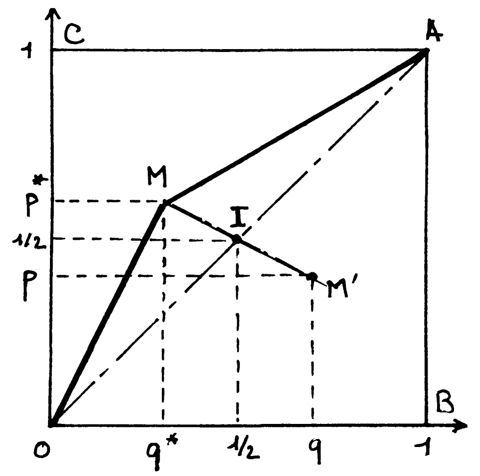
Il n'est pas toujours nécessaire de recourir à des indicateurs quantitatifs pour comparer les inégalités entre proportions. Ainsi, l'utilisation des fonctions de concentration de Gini permet de s'en dispenser dans certains cas.

8.1. Comparaison des fonctions de concentration de distributions dichotomiques

La fonction de concentration d'une distribution de fréquences dichotomiques P=(p,p*) par rapport à une distribution de fréquences dichotomiques Q=(q,q*) admet comme courbe représentative la ligne brisée (OMA) comme le montre les graphiques G6. Elle est située par convention au dessus de la première bissectrice et appartient au triangle OCA. Elle est déterminée par les coordonnées du point M : (q,p) si p≥q et (q*,p*) si p<q. La concentration est nulle si la courbe est confondue avec le segment diagonal [OA], est d'autant plus élevée que la courbe est éloignée de la première bissectrice et est maximum si la courbe est confondue avec (OCA).



Cas de figure : p ≥ q



Cas de figure : p < q

Graphiques G6 : Courbe de concentration de P par rapport à Q

De représentation simple, les fonctions de concentration de Gini permettent d'ordonner partiellement les couples de distributions que l'on compare, lorsque les courbes de concentration ne se coupent pas [3, p.615]. Dans ce cas, on peut se dispenser du recours à un indicateur quantitatif.

Les fonctions de concentration vérifient la propriété d'identité entre distributions. En effet si les distributions P et Q sont identiques, donc si p=q, la courbe de concentration correspondante est confondue avec [OA] : la concentration est nulle. Réciproquement si la courbe de concentration de P par rapport à Q est confondue avec [OA], les deux distributions sont identiques.

Par contre les fonctions de concentration ne sont pas symétriques en P et Q. Cependant, comme les courbes de concentration de Q par rapport à P et de P par rapport à Q sont symétriques par rapport à la diagonale (BC), elles conduisent au même classement des couples (P,Q). Pour effectuer des comparaisons d'inégalité entre des couples de distributions (P,Q), il suffit donc de s'appuyer soit sur les fonctions de concentration des distributions P par rapport aux distributions Q, soit sur celles des distributions Q par rapport aux distributions P.

Comparer les inégalités entre couples de distributions P=(p,p*) et Q=(q,q*) à l'aide des fonctions de concentration de P par rapport à Q revient à déterminer la position mutuelle des courbes de concentration correspondantes (OMA), donc à effectuer une double comparaison d'une part entre les pentes $\pi(P,Q)$ des droites (OM) et d'autre part entre les pentes $\pi^*(P,Q)$ des droites (MA). Ces pentes sont égales au maximum et au minimum des deux rapports de fréquences p/q et p*/q*, soit respectivement : $\pi(P,Q)=\mathrm{Max}\{p/q,p^*/q^*\}$ et $\pi^*(P,Q)=\mathrm{Min}\{p/q,p^*/q^*\}$. La concentration de P par rapport à Q s'accroît lorsqu'il y a augmentation de la pente supérieure $\pi(P,Q)$ et non augmentation de la pente inférieure $\pi^*(P,Q)$ ou diminution de la pente inférieure et non diminution de la pente supérieure. Elle diminue lorsqu'il y a diminution de la pente supérieure et non diminution de la pente inférieure ou augmentation de la pente inférieure et non augmentation de la pente supérieure. Dans ces deux cas les courbes de concentration ne se coupent pas. Par contre on ne peut pas conclure sur la variation de la concentration lorsque les deux pentes varient dans le même sens, ce qui correspond à l'intersection des courbes de concentration.

Les comparaisons fondées sur les pentes $\pi(Q,P)$ et $\pi^*(Q,P)$ des courbes de concentration de Q par rapport à P conduisent à des résultats analogues. En effet ces pentes sont liées aux pentes $\pi(P,Q)$ et $\pi^*(P,Q)$ des courbes de concentration de P par rapport à Q : $\pi(Q,P)=\mathrm{Max}\{q/p,q^*/p^*\}=1/\mathrm{Min}\{p/q,p^*/q^*\}=1/\pi^*(P,Q)$ et $\pi^*(Q,P)=\mathrm{Min}\{q/p,q^*/p^*\}=1/\mathrm{Max}\{p/q,p^*/q^*\}=1/\pi(P,Q)$. Ces relations entre pentes correspondent à la symétrie par rapport à (BC) des courbes de concentration associées.

8.2. Comparaison des inégalités sociales de scolarisation à partir des fonctions de concentration de Gini

A partir du tableau statistique de base croisant les situations scolaires et les catégories sociales, on peut composer différents couples de distributions dichotomiques susceptibles de représenter les situations d'inégalité, comme on l'a mis en évidence au paragraphe 1.3. Chacun d'eux peut donc être utilisé pour comparer les inégalités sociales de scolarisation à l'aide de fonctions de concentration de Gini.

Trois couples sont obtenus en combinant deux à deux les distributions des catégories sociales respectivement conditionnelle à l'admission (u,u*)=(ax/m,a*y/m), conditionnelle à l'exclusion (v,v*)=(ax*/m*,a*y*/m*) et marginale (a,a*), c'est à dire les distributions d'enfants de cadres et d'ouvriers respectivement parmi les admis, parmi les exclus et dans l'ensemble. Les fonctions de concentration de (u,u*) par rapport à (a,a*), de (v,v*) par rapport à (a,a*) et de (u,u*) par rapport à (v,v*) caractérisent respectivement les inégalités sociales d'admission, d'exclusion et des admissions relativement aux exclusions.

De même trois couples s'obtiennent par combinaison deux à deux des distributions de situations scolaires respectivement conditionnelle à la catégorie cadre (x,x*), conditionnelle à la catégorie ouvrier (y,y*) et marginale (m,m*)=(ax+a*y,ax*+a*y*), c'est à dire les distributions d'admis et d'exclus parmi les enfants de cadres, d'ouvriers et d'ensemble. Les fonctions de concentration de (x,x*) par rapport à (y,y*) caractérisent les inégalités de scolarisation entre enfants de cadres et d'ouvriers. Les fonctions de concentration de (x,x*) ou de (y,y*) par rapport à (m,m*) correspondent aux inégalités de scolarisation des enfants de cadres ou d'ouvriers par rapport à l'ensemble.

Pour garantir la cohérence des comparaisons dans l'acception retenue au second chapitre pour les indices d'inégalité, il convient de choisir des fonctions de concentration qui, lorsqu'on remplace les proportions x et y par les proportions complémentaires x* et y*, admettent des courbes représentatives soit identiques, soit symétriques par rapport à (BC). Tel est le cas des fonctions de concentration fondées sur les distributions de situations scolaires ou celle de (u,u*) par rapport à (v,v*). Par contre cette cohérence n'est pas obtenue avec les fonctions de concentration de (u,u*) ou de (v,v*) par rapport à (a,a*). On retrouve ainsi dans le cas de distributions dichotomiques, la distinction établie pour des distributions quelconques, entre concentration portant sur les admis et concentration portant sur les exclus [4, pp.14-15; 5, pp.11-17].

Les pentes π et π^* qui permettent de comparer les fonctions de concentration associées à ces couples de distributions ont des expressions différentes et conduisent à des résultats divergents, comme le montre le tableau T4. Avec les données de l'exemple retenu, les fonctions de concentration ne sont suffisantes pour comparer les inégalités que pour deux couples de distributions. On observe une augmentation de la concentration de (v,v*) par rapport à (a,a*), donc des inégalités sociales d'exclusion. On constate également une augmentation de la concentration de (x,x*) par rapport à (m,m*), donc des inégalités de scolarisation des enfants de cadres par rapport à l'ensemble. Mais on ne peut pas conclure dans les autres cas, notamment sur l'évolution de la concentration de (u,u*) par rapport à (a,a*) caractérisant les inégalités sociales d'admission, ou sur celle de la concentration de (x,x*) par rapport à (y,y*) caractérisant les inégalités de scolarisation entre enfants de cadres et d'ouvriers.

Distributions	Formulation des pentes	Pente supérieure		Pente inférieure		Evol. concent
		en t_1	en t_2	en t_1	en t_2	
(u,u*) / (a,a*)	x/m et y/m	2,06	2,38	0,06	0,38	?
(v,v*) / (a,a*)	y*/m* et x*/m*	1,21	1,22	0,77	0,51	+
(u,u*) / (v,v*)	m*x/mx* et m*y/my*	2,68	4,64	0,05	0,31	?
(x,x*) / (m,m*)	x/m et x*/m*	2,06	2,38	0,77	0,51	+
(y,y*) / (m,m*)	y*/m* et y/m	1,21	1,22	0,06	0,38	?
(x,x*) / (y,y*)	x/y et x*/y*	37	6,2	0,64	0,42	?

Tableau T4
Comparaison des fonctions de concentration de Gini

8.3. Indices d'inégalité définis à partir des indices de concentration de Gini

Lorsque les courbes de concentration sont concourantes, on est amené à utiliser des indicateurs quantitatifs pour comparer les inégalités entre distributions. Parmi eux, outre les indices présentés dans les chapitres précédents, figurent les *indices de concentration de Gini G* associés aux fonctions de concentration. Ils ont pour valeur le double de l'aire comprise entre la courbe de concentration et la première bissectrice. L'indice de concentration d'une distribution P=(p,p*) par rapport à une distribution Q=(q,q*) a pour expression algébrique $G(P,Q)=|p-q|$. Symétrique en p et q, il est indépendant de la distribution choisie comme référence : $G(Q,P)=G(P,Q)$.

Appliqués aux différents couples de distributions dichotomiques issus du tableau statistique de base, les indices de concentration présentent des formulations très diverses, comme le montre le tableau T5, et interdépendantes puisque $G3_a=G1_a+G2_a$ et $G6=G4_a+G5_a$. Compte tenu des relations entre proportions et fréquences, ils peuvent tous s'exprimer en fonction des proportions x et y et de la fréquence a. Ils constituent donc des familles d'indices de paramètre a.

Distributions	Formulation	Valeur en t_1	Valeur en t_2	Evolution
(u,u*) et (a,a*)	$G1a(x,y)=aa^*\|x-y\|/m$	0,50	0,43	-
(v,v*) et (a,a*)	$G2a(x,y)=aa^*\|x-y\|/m^*$	0,11	0,15	+
(u,u*) et (v,v*)	$G3a(x,y)=aa^*\|x-y\|/mm^*$	0,61	0,58	-
(x,x*) et (m,m*)	$G4a(x,y)=a^*\|x-y\|$	0,19	0,36	+
(y,y*) et (m,m*)	$G5a(x,y)=a\|x-y\|$	0,17	0,16	-
(x,x*) et (y,y*)	$G6(x,y)=\|x-y\|$	0,36	0,52	+

Tableau T5
Formulation et évolution des indices de concentration de Gini

L'indice G6 et les indices des familles ($G4_a$) et ($G5_a$) définis à partir des distributions des situations scolaires sont proportionnels à l'écart absolu, avec un coefficient de proportionnalité ne dépendant que de a : G6=E; $G4_a=a^*E$; et $G5_a=aE$. Egal à l'écart absolu, G6 est un indice d'inégalité cohérent. Par contre $G4_a$ et $G5_a$ ne sont pas des indices d'inégalité car ils ne sont pas symétriques en x et y : $G4_a^*(y,x)=G5_a(x,y)$ et $G5_a^*(y,x)=G4_a(x,y)$.

Les indices des familles ($G1_a$), ($G2_a$) et (G3v) définis à partir des distributions des catégories sociales suivent les propriétés de positivité, d'identité et de symétrie, comme le montre leur formulation du tableau T5. Comme ils respectent également la propriété de variation en V sur $]0,1[^2$, ils constituent des familles d'indices d'inégalité.
Démonstration :

$G1_a(x,y)=aa^*|x-y|/m=aa^*|x-y|/(ax+a^*y)$

Soit $A_a(x,y)=(x-y)/(ax+a^*y)$

$G1_a(x,y)=aa^*A_a(x,y)$ si $x>y$; 0 si $x=y$ et $-aa^*A_a(x,y)$ si $x<y$.

$\forall y>0 \quad \partial A_a/\partial x(x,y)=y/(ax+a^*y)^2>0$

d'où : $\partial G1_a/\partial x(x,y)>0$ si $x>y$ et $\partial G1_a/\partial x(x,y)<0$ si $x<y$

$\forall (x,y)\in]0,1[^2$ tel que $x\neq y$ $(x-y)\partial G1_a/\partial x(x,y)>0$

On démontre de même que : $\forall (x,y)\in]0,1[^2$ tel que $x\neq y$ $(y-x)\partial G1_a/\partial y(x,y)>0$

$G1_a$ respecte la propriété de variation en V sur $]0,1[^2$.

On le vérifie de la même façon pour $G2_a$. En tant que somme de $G1_a$ et de $G2_a$, $G3_a$ suit également cette propriété.

La propriété de cohérence est vérifiée par $G3_a$, mais pas par $G1_a$ et $G2_a$: $G1_a(x^*,y^*)=G2_a(x,y)$ et $G2_a(x^*,y^*)=G1_a(x,y)$. On aboutit donc à un résultat analogue à celui qui a été établi avec les fonctions de concentration. En outre la propriété d'homogénéité en x et y est suivie par $G1_a$.

En définitive, parmi les indices de concentration construits à partir du tableau statistique de base, seuls G6 et ceux de la famille (G3a) constituent des indices d'inégalité cohérents.

Avec les données de l'exemple retenu, les indices de concentration conduisent à des résultats différents selon les distributions sur lesquelles ils portent (tableau T5). En particulier G6 montre une augmentation des inégalités de scolarisation entre enfants de cadres et d'ouvriers et $G3_a$ une diminution. Par ailleurs on constate une diminution de G1a et une augmentation de $G2_a$. On aboutit donc à une évaluation contradictoire de l'évolution des inégalités sociales de scolarisation, selon qu'elle se fait en référence à l'admission ou à l'exclusion. Un tel résultat illustre les problèmes d'interprétation qui peuvent surgir dans les comparaisons reposant sur des indices non cohérents. On vérifie par ailleurs que les indices de concentration aboutissent à des résultats analogues à ceux qui ont été observés lorsque les comparaisons entre fonctions de concentration sont concluantes. Ainsi aux augmentations des concentrations de (v,v*) par rapport à (a,a*) et de (x,x*) par rapport à (m,m*) correspondent celles des indices $G2_a$ et $G4_a$.

9. CONCLUSION

Les indicateurs d'inégalité entre proportions présentent une grande diversité. Il en résulte qu'ils conduisent souvent à des résultats contradictoires lorsqu'ils sont utilisés pour des comparaisons. On le vérifie notamment avec l'exemple de l'évolution de l'inégalité sociale d'accès à l'enseignement secondaire britannique au cours de la première moitié du siècle.

Les indicateurs d'inégalité relèvent de différents modes de construction. Certains, comme ceux qui se définissent à partir des fonctions de concentration de Gini, ressortissent à des méthodes non quantitatives de comparaison. Ces indicateurs de concentration se distinguent, et il convient de les choisir en conséquence, selon les couples de distributions sur lesquelles ils peuvent porter et que l'on peut composer à partir des données des tableaux statistiques de base croisant par exemple situations scolaires et catégories sociales. Les comparaisons fondées sur les fonctions de concentration n'aboutissent pas toujours à des résultats concluants. Il convient alors d'utiliser des indicateurs quantitatifs. Nous avons étudié en particulier ceux qui se définissent à partir de mesures, de distances, d'écarts moyens, des indices d'Atkinson ou de l'indice de concentration de Gini. Ces différents indicateurs ne sont pas exclusifs les uns des autres, comme le montrent les relations qui s'établissent entre eux.

La diversité des indicateurs d'inégalité quantitatifs reflète des différences de propriétés et de formes, dont il convient de tenir compte dans les choix. Il convient tout d'abord de vérifier que les indicateurs respectent les propriétés de positivité, d'identité, de symétrie et de variation en V qui les caractérisent comme indices d'inégalité entre proportions. Il convient ensuite de fonder leur choix en fonction d'un certain nombre de propriétés alternatives qu'il peut être souhaitable de leur attribuer : cohérence ou homogénéité; relativité ou non par rapport à une proportion de référence; pondération ou non des catégories comparées. Il convient enfin de choisir les caractéristiques de forme des indices en tenant compte des présupposés théoriques adoptés pour interpréter les inégalités. Il s'agit par exemple de déterminer la forme des densités μ des indices associés à des mesures, ou celle des fonctions f des indices définis à partir de distances fonctionnelles ou d'écarts fonctionnels moyens. De ce point de vue il est intéressant d'utiliser des indices paramétriques comme ceux des familles cohérentes (Φ_α), $(\Psi_{r\alpha})$, $(E_{\alpha rpa})$ ou de la famille homogène (S_α), dont les paramètres permettent de représenter les inégalités sous une grande variété de formes.

Les indicateurs d'inégalité entre proportions étudiés dans cet article portent sur deux catégories. Ce sont des indicateurs synthétiques lorsque ces catégories sont dichotomiques, comme dans le cas des comparaisons d'inégalité de scolarisation entre garçons et filles. Mais lorsque ces catégories ne couvrent pas l'ensemble des modalités possibles, comme dans le cas des catégories cadre et ouvrier dans la population active, ce sont des indicateurs d'inégalité partiels. L'utilisation de tels indicateurs s'impose lorsqu'on focalise les comparaisons sur deux catégories particulières, par exemple si l'on s'intéresse aux inégalités des situations scolaires entre les catégories cadre et ouvrier. Mais si l'on veut comparer les inégalités sociales d'ensemble vis à vis d'une variable, comme par exemple la proportion d'admis, à partir des données de deux catégories sociales contrastées sélectionnées parmi les n (n>2) catégories de la population active, il convient de les choisir suffisamment représentatives des inégalités d'ensemble. Et lorsque les données sont disponibles, il convient de compléter ces comparaisons par celles qui se fondent sur des indices synthétiques d'inégalité portant sur l'ensemble des catégories[11] [3, pp.617].

---

[11] La question des indices synthétiques d'inégalité est abordée dans un autre article de l'auteur [18]. L'étude est effectuée dans le cas de distributions interrégionales, mais les résultats s'appliquent à toute autre forme de distribution, notamment aux distributions entre catégories sociales dont il est question ici.

BIBLIOGRAPHIE

[1] ATKINSON A. B., "On the measurement of inequality", *Journal of Economic Theory*, 2 (1970), 244-263.

[2] BACCINI A., de FALGUEROLLES A., QANNARI M., "Etude de quelques indices de concentration : un essai de présentation unifiée", *Revue de Statistiques Appliquées*, 34 (1986), 31-44.

[3] BARBUT M., "Note sur quelques indicateurs globaux de l'inégalité : C. Gini, V. Pareto, P. Lévy", *Revue Française de Sociologie*, XXV (1984), 609-622.

[4] BARBUT M., "Sur quelques propriétés élémentaires des fonctions de concentration de C. Gini", *Mathématiques et Sciences humaines* , 88 (1984), 5-20.

[5] BARBUT M., "Sur les indicateurs de l'inégalité : croissance logistique et mesure de l'inégalité, et de quelques effets paradoxaux dans la comparaison des inégalités", *Mathématiques et Sciences humaines* , 90 (1985), 5-17.

[6] BOUDON R., *L'inégalité des chances*, Paris, A. Colin, 1973.

[7] BOURGUIGNON F., "Decomposable income inequality measures", *Econometrica*, 47 (1979), 901-920.

[8] COMBESSIE J.-C., "L'évolution comparée des inégalités : problèmes statistiques", *Revue Française de Sociologie*, XXV (1984), 233-254.

[9] COMBESSIE J.-C., "Paradoxes des fonctions de concentration de C. Gini", *Revue Française de Sociologie*, XXVI (1985), 653-658.

[10] FLORENS J.-P., "Inégalité et dépendance statistique", *Revue Française de Sociologie*, XXV (1984), 255-263.

[11] GRÉMY J.-P., *Introduction à la lecture des tableaux statistiques*, Laboratoire d'étude des méthodes et techniques de l'analyse sociologique, Paris, multigraphié, 1977.

[12] GRÉMY J.-P., "Sur les différences entre pourcentages et leur interprétation", *Revue Française de Sociologie*, XXV (1984), 396-420.

[13] GRÉMY J.-P., "Sur l'utilisation d'indices numériques pour mesurer l'inégalité", *Mathématiques et Sciences humaines* , 93 (1986), 7-40.

[14] MERLLIÉ D., "Analyses de l'interaction entre variables. Problèmes statistiques ou sociologiques?", *Revue Française de Sociologie*, XXVI (1985), 629-652.

[15] PRÉVOT J., "A propos d'indices et de comparaisons de proportions", *Revue Française de Sociologie*, XXVI (1985), 601-628.

[16] SHORROCKS A.-F., "The class of additively decomposable inequality measures", *Econometrica*, 48 (1980), 613-625.

[17] THEIL H., *Economics and Information Theory*, North-Holland, Amsterdam, 1967.

[18] VALEYRE A., "Mesures de dissemblance et d'inégalité interrégionales : principes, formes et propriétés", *Revue d'Economie Régionale et Urbaine*, (1993), 17-53.

[19] VALEYRE A., *Mesures d'inégalité entre proportions*, Laboratoire Techniques, Territoires et Sociétés, Noisy-le-Grand, multigraphié, 1994.

[20] VALLET L.-A., "L'évolution de l'inégalité des chances devant l'enseignement. Un point de vue de modélisation statistique", *Revue Française de Sociologie*, XXIX (1988), 395-423.

[21] WESTERGAARD J., LITTLE A., "The trend of class differentials in educational opportunity in England and Wales", *British Journal of Sociology*, XV (1964), 301-316.