Groupe de Recherche en Économie et Développement International

Cahier de recherche / Working Paper
06-11

La décomposition de l'indicateur de Gini en sous-groupes : une revue de la littérature

Stéphane Mussard

Françoise Seyte

Michel Terraza

UNIVERSITÉ DE SHERBROOKE

La décomposition de l'indicateur de Gini en sous-groupes : une revue de la littérature

Stéphane Mussard[☯]

Françoise Seyte[*][1]

Michel Terraza[*]

[☯] GREDI, Département d'économique,
Université de Sherbrooke
GEREM, Département des sciences économiques
Université de Perpignan

[*] LAMETA, Université de Montpellier I,
UFR Sciences Economiques,
Avenue de la Mer, Site de Richter - CS 79606
34960 Montpellier Cedex 2
France

[1] Auteur correspondant : f-seyte@lameta.univ-montp1.fr

Les inégalités économiques font partie de l'économie publique et du domaine de la justice sociale. Pour apprécier l'efficacité d'un principe de justice sociale (commutative, distributive, locale, etc.), les mesures d'inégalité apparaissent comme un outil privilégié puisqu'il est possible de les utiliser avant et après l'application d'une mesure de justice sociale (ou de politique économique redistributive). Les indicateurs d'inégalité forment donc un préalable aux prises de décisions publiques concernant les politiques socio-économiques.

Cet article s'intéresse uniquement aux aspects concernant les mesures de l'inégalité du revenu. Cette partie de l'économie publique est issue de la théorie de la répartition du revenu qui comprend deux parties distinctes : la répartition fonctionnelle du revenu et la répartition personnelle du revenu.[1]

La répartition fonctionnelle du revenu étudie la formation des prix, le rôle des facteurs de production dans la production totale et l'allocation du revenu aux propriétaires de chaque catégorie de facteurs. Elle permet donc de déterminer la part du revenu national imputée à chaque catégorie de facteurs de production.

La répartition personnelle du revenu analyse la répartition du revenu total entre les unités économiques (famille, ménage, individus, etc.). Cet axe de recherche se focalise aussi sur les sources de revenu de chaque entité, telles que la masse du salaire ou la masse du profit. Ce domaine repose sur la propriété privée des ressources initiales.

La démarche que nous adoptons appartient au domaine de la répartition personnelle du revenu. Plus précisément, il s'agit de l'approche par les indices de mesure des inégalités, notamment impulsée par les travaux de Kolm (1966), Atkinson (1970) et Sen (1973). Ce domaine s'attache à mesurer l'intensité des inégalités de revenu entre les individus d'une même population afin de savoir si ces disparités sont jugées importantes ou non. La valeur de l'indice indique donc le caractère égalitaire ou inégalitaire de la répartition des revenus.

L'approche par les indices est différente de celle par les modèles de distributions de revenu, qui tente à partir d'estimation économique de s'adapter aux revenus observés. L'approche par les distributions autorise l'évaluation des inégalités en interprétant l'intensité et l'évolution des paramètres d'un modèle estimé. Il s'agit d'estimations paramétriques des disparités de revenu. En ce domaine, Pareto (1896) fait figure de précurseur : il est l'un des premiers à s'être intéressé à la modélisation de la

[2] Ces deux types de répartitions ne sont pas totalement cloisonnés. Confer les travaux de Dagum (1999) sur le lien qui prévaut entre la répartition personnelle et la répartition fonctionnelle du revenu.

répartition personnelle du revenu. On recense par la suite les modèles de Champernowne (1952) ou encore de Singh et Maddala (1976). Les modèles de C. Dagum (1977) constituent de nos jours les modèles les plus robustes dont la spécification, issue des travaux de Pareto, revêt une cohérence à la fois économique et économétrique.

De plus en plus, ces méthodes paramétriques ont été accompagnées d'estimations non paramétriques portant sur les indices d'inégalité. Ces indices, simples d'interprétation, car souvent compris entre zéro et un, autorisent des comparaisons inter populations. L'estimation par les indices permet d'expliquer la totalité des inégalités de revenu observées, sans perte d'information. Il est cependant nécessaire de vérifier les différentes propriétés qu'ils intègrent.

La décomposition en sous-groupes des indices d'inégalité est une propriété fondamentale. Les décompositions des mesures d'inégalité du revenu offrent des possibilités de comparaison intéressantes. Elles permettent de cibler les groupes qui tendent à accroître les inégalités d'une société. Cependant, toutes les mesures d'inégalité ne sont pas munies de cette propriété. Les fondements même de cette décomposition reposent sur des aspects mathématiques précis, comme la séparabilité additive d'une fonction ce qui nous conduit à étudier les propriétés issues des mesures décomposées, et en particulier celles du coefficient de Gini (1) avant de recenser de manière non exhaustive ses principales décompositions (2).

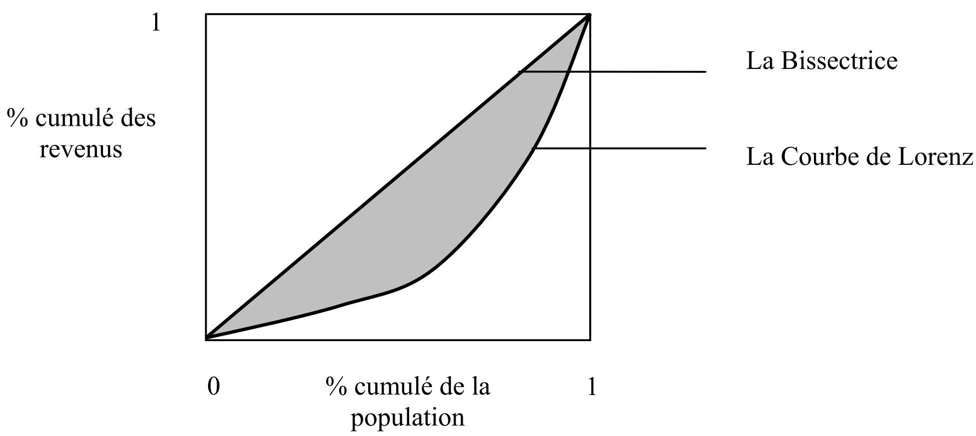
I. la mesure des inégalités de revenu :

I.1. Les mesures de Gini et de l'entropie :

I.1.1. Le coefficient de Gini :

L'indice de Gini (1921) se calcule à partir de la courbe de Lorenz (1905) qui relie les proportions cumulées de la population en classes (centiles, déciles, …) avec les pourcentages cumulés des revenus correspondants. Il représente deux fois l'aire contenue entre la première bissectrice et la courbe de Lorenz et il est compris dans l'intervalle [0,1]. Il mesure les inégalités de répartition de revenu et la concentration d'une distribution. Plus l'indice tend vers 1, plus la répartition des revenus est inégalitaire et la concentration forte. En revanche, si l'indicateur est égal à 0, la répartition des revenus est égalitaire et la concentration faible. L'indice de Gini représente aussi (en le multipliant par deux et par la moyenne de la distribution) l'écart de revenu espéré entre deux individus tirés au hasard dans une population (avec remise) [confer Pyatt (1976), p. 244].

Figure 1 : Le coefficient de Gini



Propriétés :

L'indicateur de Gini respecte le principe de normalisation. Cette propriété autorise la comparaison des indices dans le temps et entre des groupes différents. L'indice permet d'étudier l'impact des transferts de revenu, notamment des riches vers les pauvres. La structure de l'indicateur de Gini est, par conséquent, liée aux politiques de redistribution qu'il est possible d'effectuer au sein d'une population afin de parvenir à un meilleur niveau d'équité.

Avantages et Inconvénients :

La mesure de Gini étant comprise entre 0 et 1, il existe une infinité de distributions de revenu différentes pour lesquelles l'indice possède la même valeur. Ainsi, pour deux populations différentes, on peut estimer un indice identique. Pour remédier à ce problème, on analyse la structure des courbes de Lorenz, notamment leur forme, par la méthode des coudes, afin de distinguer des différences au niveau des groupes riches, des groupes pauvres, et de la classe moyenne [cf. Mesnard (1997)]. La modélisation des distributions de revenu par les modèles de Dagum (1977) permet la comparaison de deux distributions, même si les courbes de Lorenz sont sécantes, car les paramètres du modèle (1 paramètre d'échelle, 1 paramètre d'inégalité, et 2 paramètres d'égalité) sont uniques pour chaque courbe et les indices de Gini peuvent être interprétés ; d'où la relation biunivoque entre les courbes de Lorenz et les modèles de Dagum.

L'indicateur de Gini possède donc des propriétés intéressantes. Des applications se sont même développées, en dehors de la mesure des inégalités de revenu, puisqu'il peut servir à apprécier les disparités dans d'autres domaines, comme la santé ou la malnutrition [voir Wagstaff (2003)] et de manière plus générale au sein de l'économie du développement.

I.1.2. L'approche de Theil (1967) et la mesure de l'entropie généralisée :

La mesure de l'entropie généralisée, dérivée des mesures thermodynamiques de l'entropie, provient des travaux de Theil (1967).[2] Il déduit un indicateur d'inégalité en s'appuyant sur la théorie de l'information. On considère un espace où chaque événement est associé à une certaine probabilité. Une fonction d'information f permettant de mesurer la probabilité de ces événements est construite sur la base de trois axiomes [confer Cowell (2000)].

(i) Si l'événement est certain, la probabilité est égale à 1 (p = 1). Dans ce cas la fonction d'information est nulle [$f(1) = 0$] car l'événement ne donne aucune information.

(ii) Les fortes probabilités ont des faibles valeurs : $p > p_1 \Rightarrow f(p) < f(p_1)$.

(iii) La valeur de l'information issue de deux événements indépendants est la somme des valeurs d'information de chaque événement : $f(pp_1) = f(p) + f(p_1)$.

La forme fonctionnelle de ce soubassement axiomatique peut être définie par la fonction logarithmique qui satisfait ces propriétés : $f(p) = -\log(p)$. L'entropie est l'information espérée dans une distribution.[3] Elle mesure aussi le désordre d'un système thermodynamique. En reliant le concept de désordre au concept d'inégalité et en remplaçant les probabilités par les parts de revenu, Theil introduit sa propre mesure d'inégalité :

$$S_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{\mu}\log\frac{x_i}{\mu}$$

[2] Confer les travaux de Foster (1983) pour une caractérisation de l'indicateur de Theil et les travaux de Hart (1970) sur les différentes mesures de l'entropie.

où x_i est le revenu de l'individu i, et μ la moyenne de la distribution (composée de n individus). Cet indicateur appartient à la classe de l'entropie généralisée définie par :

$$Sc = \sum_{i=1}^{n} \frac{\left((x_i/\mu)^c - 1\right)}{nc\,(c-1)} \;, \forall c \neq 0{,}1\,,$$

où le paramètre $c \in\ ]-\infty,+\infty[$ mesure la sensibilité de l'indice par rapport aux différentes parties de la distribution. Des indices sensibles aux changements dans les hauts revenus (partie supérieure de la distribution) ont des paramètres positifs et élevés. Des paramètres négatifs renvoient à des indicateurs sensibles aux changements dans la partie inférieure de la distribution. Si c = 1, on atteint en limite la mesure de Theil (S_1). Si c = 0, on obtient l'écart moyen des logarithmes (représenté par le sigle MLD ou par S_0):

$$S_0 = \frac{1}{n}\sum_{i=1}^{n} \log\frac{\mu}{x_i}\,.$$

Si c = 2, l'entropie généralisée se définit comme la moitié du coefficient de variation au carré (S_2) :

$$S_2 = \frac{\sigma^2}{2\mu^2} = \frac{1}{2}V^2(x),$$

où σ est l'écart-type et V le coefficient de variation : σ/μ.

Les mesures d'inégalité de Gini et celles issues de l'entropie sont les mesures les plus connues et les plus utilisées. Ces indicateurs possèdent des propriétés spécifiques. C'est pourquoi, il est nécessaire de faire appel aux différentes approches de l'économie des inégalités pour évaluer les implications socio-économiques de ces différentes propriétés, étape nécessaire avant de s'intéresser à la décomposition de la mesure.

I.2. Les différentes approches des mesures d'inégalité :

L'économie des inégalités réunit trois approches : l'approche positive, l'approche normative et l'approche axiomatique.[4] L'approche positive, « ce qui est », permet l'évaluation empirique des disparités. Le domaine normatif montre qu'il existe des relations entre les fonctions de bien-être social et les mesures d'inégalité du revenu. L'approche axiomatique tente de répertorier, comparer, et généraliser certaines classes de mesures d'inégalité.

[3] Pour plus d'explications sur l'entropie et les mesures dérivées, confer Hart (1970) et Dagum (1997b).
[4] Il est possible de rajouter à ces trois approches : l'approche expérimentale et l'approche subjective.

I.2.1. L'approche axiomatique :

L'approche axiomatique des mesures d'inégalité s'est notamment développée sous l'impulsion des travaux de Dalton (1920), Kolm (1976a, 1976b), Shorrocks (1980) et Chakravarty (1999)[5]. Elle se base sur la formulation de propriétés mathématiques permettant d'avoir un ensemble d'indices pour établir des choix entre les distributions de revenu deux à deux[6]. L'analyse axiomatique constitue un moyen de sélection des mesures d'inégalité. On retiendra les indices possédant un maximum de propriétés tout en vérifiant la compatibilité de ces principes entre eux.

Enfin, cette approche permet d'établir de nouvelles classes de mesures en synthétisant plusieurs axiomes. Par exemple, la classe des indices centristes [Kolm (1976a, 1976b)] est issue d'un compromis entre l'ensemble des mesures d'inégalité absolues et relatives [voir aussi Bossert et Pfingsten (1990)]. Les principaux axiomes respectés par les mesures d'inégalité sont les suivants :

→ La continuité

L'indice d'inégalité est une fonction continue :

$$I : \Re^n \rightarrow \Re_+ \ \forall n \in \mathrm{IN}, \qquad \text{(CN)}$$

où $\Re^n$, $\Re_+$ et IN sont respectivement l'ensemble des réels de dimension n, l'ensemble des réels non négatifs et l'ensemble des entiers naturels.

→ Le principe de transfert de Pigou-Dalton

Un transfert engendre la baisse d'un indice (diminution des inégalités au sein de la société) lorsqu'un individu riche reverse une partie de son revenu à une personne moins riche que lui.[7] Inversement, l'indice augmente (caractérisant une augmentation des inégalités au sein de la société) quand un transfert est pratiqué d'une personne pauvre vers une personne riche. Si une distribution x est obtenue à partir d'une distribution y où l'on effectue un transfert de revenu progressif (d'une personne riche vers une personne

[5] Les travaux axiomatiques de Chakravarty sont antérieurs à 1999, mais l'article de 1999 du Handbook of Income Inequality Measurement offre une revue de littérature très appréciable.

[6] Une propriété axiomatique est un ensemble de conditions susceptibles d'être vérifiées par l'indicateur d'inégalité. Par exemple, on utilise des indices normalisés, c'est-à-dire des indices compris dans l'intervalle [0,1], autorisant la comparabilité des valeurs.

[7] Les transferts sont valables que les individus conservent ou non leur rang dans la distribution.

pauvre) alors la mesure des inégalités associée à x devient plus faible que celle associée à y :

$$I(x) < I(y). \tag{PD}$$

→ Le principe de population de Dalton

Le principe de population indique qu'un indice reste inchangé lorsque la population s'accroît de manière identique. Si chaque individu d'une population se retrouve avec une personne ayant le même revenu, les inégalités restent inchangées. Pour une distribution $x = \{x_1, x_2, \ldots, x_n\}$ que l'on réplique k fois :

$x = \{\overbrace{x_1,\ldots,x_1}^{k \text{ fois}},\ldots,\overbrace{x_n,\ldots,x_n}^{k \text{ fois}}\}$, l'indice d'inégalité reste inchangé,

$$I^k(x) = I(x), \forall\, k \geq 2. \tag{PP}$$

→ La symétrie (Anonymat)

Cette propriété montre qu'un indice est invariant lorsque le rang d'un individu dans la distribution est modifié. Par exemple, si l'on modifie l'ordre dans la distribution (en classant les revenus par ordre croissant ou décroissant), la mesure d'inégalité doit conserver la même valeur. Ce principe désigne la propriété d'un indice invariant après permutation de l'ordre des individus ou permettant à chacun de conserver son anonymat. Si x est obtenu à partir de y par permutation des revenus, on obtient alors une fonction symétrique de ses arguments :

$$I(x) = I(y). \tag{SM}$$

→ La normalisation

Elle autorise la comparaison d'indices d'inégalité compris dans l'intervalle [0,1]. Si l'indice tend vers 0 les revenus sont répartis de manière égalitaire ; en revanche, si l'indicateur tend vers 1 la répartition est jugée inégalitaire. On associe la normalisation à la condition suivante :

$$I(k1^n) = 0, \forall\, k > 0, \tag{NM}$$

où 1^n est le vecteur unité de dimension n (vecteur dont les éléments sont tous égaux à 1). Il s'agit de la condition selon laquelle les indices d'inégalité prennent la valeur 0 pour des distributions égalitaires.

L'ensemble formé par les mesures qui satisfont les axiomes (CN), (PD), (PP), (SM) et (NM) est appelé ensemble des indices réguliers. Si on ajoute la propriété d'invariance relative, on définit la classe des indices réguliers relatifs.

→ L'invariance relative

La propriété d'invariance relative stipule que lorsqu'une population voit ses revenus multipliés par un nombre supérieur à zéro (par exemple multipliés par deux) alors les inégalités restent inchangées[8]. La propriété d'invariance relative se traduit par une fonction homogène de degré zéro. Elle s'exprime sous la forme :

$$I(\lambda x) = I(x), \ \forall \ \lambda > 0. \qquad \text{(IR)}$$

Les mesures relatives ne sont pas affectées par une augmentation proportionnelle des revenus. Si les inégalités sont mesurées, par exemple, sur des revenus exprimés en euros ou bien sur ces mêmes revenus exprimés en dollars, l'indice d'inégalité procure le même résultat.

I.2.2. L'approche normative :

L'approche normative dicte « ce qui doit être ». Elle se base sur des propriétés reflétant des valeurs éthiques qui doivent être respectées par les indicateurs. Kolm (1977) spécifie clairement l'existence d'une relation duale entre les fonctions de bien-être social représentant les préférences des individus d'une même société, et les mesures d'inégalité du revenu. Ces travaux seront ensuite poursuivis et étendus par Blackorby, Bossert et Donaldson (1999).[9] Dagum (1998) démontre aussi le lien entre les fonctions de bien-être social et les mesures d'inégalité. Il introduit trois principes de base :

***P*1.** Les agents ont la possibilité de faire des comparaisons interpersonnelles d'utilité. Les individus se livrent sans cesse à des comparaisons entre la satisfaction qu'ils retirent d'une situation (par exemple un bien de consommation) et celle que les autres individus ressentent. Ce principe est fondamental. En effet, selon Dagum, les processus économiques s'insèrent à l'intérieur des sociétés ; les unités et sous-

[8] Cette propriété suscite un intérêt particulier et controversé, car certaines mesures ne satisfont pas l'invariance relative mais l'invariance dite absolue : les inégalités de revenu ne sont pas affectées lorsqu'on octroie la même somme à chaque individu. La littérature montre que l'on abandonne progressivement ces deux axiomes pour des concepts intermédiaires.

[9] Pour une explication de l'approche normative confer Blackorby, Bossert et Donaldson (1999).

ensembles d'unités économiques font continuellement des comparaisons interpersonnelles entre leurs niveaux et leurs différences d'utilité. Dagum (1998) spécifie ces fonctions en les dérivant des mesures d'inégalité du revenu. Le principe *P*1 autorise la définition d'une fonction d'utilité pour chaque individu (fonction mathématique représentant la satisfaction) déterminée à la fois par le revenu de l'individu et les revenus des autres individus.

***P*2.** Les individus ont une préférence pour l'accumulation de richesse. Lorsque leur revenu croît, ils retirent de plus en plus de satisfaction de leur situation, avec cependant une augmentation de satisfaction liée à la dernière unité de richesse acquise de moins en moins importante. Le principe *P*2 désigne la détermination des agents économiques à offrir notamment leur force de travail (c'est-à-dire des facteurs de productivité marginale) afin d'acquérir un revenu croissant pour intensifier leur satisfaction.

***P*3.** Les individus ont une préférence pour l'équité ou une aversion à l'inégalité. Cette propriété suppose que les individus d'une société préfèrent des situations où les revenus sont répartis de manière égalitaire à des situations où la société est composée d'individus dont les revenus sont répartis de manière inégalitaire[10]. La propriété *P*3 désigne la volonté de toute société à atteindre une plus grande équité, mais selon Dagum la parfaite égalité dans la répartition du revenu (utilisée dans la spécification de certains indices) n'a jamais été le but ni une condition de maximisation du bien-être social, même pour des populations d'individus identiques.

Les approches axiomatiques et normatives permettent de mettre en évidence des ensembles de propriété vérifiés par les indices d'inégalité. La littérature montre que seule, la mesure de Gini satisfait les axiomes et les propriétés mentionnées ci-dessus. En effet, l'entropie généralisée respecte les cinq axiomes définissant l'ensemble des mesures régulières et relatives, mais Dagum (1998) montre qu'elle ne satisfait pas *P*1. Par conséquent, étant donné que *P*1 est un principe fondamental, l'indicateur de Gini peut être privilégié à la mesure de l'entropie généralisée ; c'est pourquoi, il constitue une des mesures les plus utilisées. Il est néanmoins possible de se demander si les trois principes coïncident entre eux [cf. Mussard, Seyte, Terraza (2003)].Les auteurs arrivent

[10] Le caractère d'une situation égalitaire et d'une situation juste est différent. En effet, une population définie par des revenus répartis de manière inégalitaire peut être jugée juste.

à la même conclusion que Dagum car la spécification statistique de la mesure de Gini est plus complète, notamment dans la décomposition en sous-groupes [cf. Dagum (1997a, 1997b)], malgré quelques oppositions [cf. Mookherjee et Shorrocks (1982) ou Salas (2002)].

La notion de décomposabilité s'inscrit dans l'approche axiomatique (confer le paragraphe suivant). Cependant les indicateurs d'inégalité reflètent des normes particulières, comme celles présentées ci-dessus. Par conséquent, l'approche par les indices décomposables est connexe à l'approche axiomatique, normative, et positive

I.3. La décomposition d'une mesure d'inégalité **:**

Pour décomposer un indicateur en sous-groupes, il est nécessaire que la population globale *P* soit divisée en plusieurs groupes (par exemple hommes et femmes, catégories socioprofessionnelles, régions, groupes d'âge, etc.). Une décomposition en sous-populations explique les inégalités de revenu par le degré d'implication des différents groupes composant *P*. Un groupe participe-t-il plus qu'un autre à l'explication des inégalités ? Il s'agit là de la première question à laquelle la décomposition peut répondre. Mais pourquoi peut-on privilégier ces approches multivariées aux approches standards non décomposées ?

L'approche non décomposée permet de comparer différents indices, s'ils sont normalisés. Cette comparaison autorise un classement des différentes sous-populations en spécifiant les inégalités qui prévalent à l'intérieur de celles-ci. La décomposition permet de surcroît de définir les groupes qui possèdent les plus fortes contributions à l'explication de l'inégalité totale. Le procédé de décomposition autorise l'estimation des inégalités à l'intérieur de chaque groupe (mesures intragroupes) et des inégalités entre les différents groupes (mesures intergroupes). Les questions auxquelles les décompositions peuvent répondre sont les suivantes : quelle est la part d'inégalité intergroupe dans l'inégalité totale, quelle est la contribution des inégalités intragroupes à l'inégalité totale, quelle est la participation d'un groupe particulier à l'inégalité totale ou encore quelle est la contribution des inégalités entre deux groupes à l'inégalité totale ?

Parmi les propriétés axiomatiques (CN), (PD), (PP), (SM), (NM) et (IR), le respect du critère de transfert est important car ses implications sont fondamentales au niveau des politiques socio-économiques de redistribution. Mais depuis les travaux de Bourguignon (1979), Cowell (1980a, 1980b), Shorrocks (1980, 1984, 1988) et Ebert

(1988) la décomposabilité est devenue tout aussi primordiale. Elle permet en effet de préciser les déterminants des inégalités au sein d'une population – étape primordiale avant d'effectuer des transferts redistributifs.

I.3.1. Les mesures agrégatives et additivement décomposables : Bourguignon (1979)

Les premières recherches sur les mesures d'inégalité décomposables sont des tentatives d'axiomatisation qui essaient de dériver ou de créer certains indices. Bourguignon (1979) est le premier à proposer des propriétés de généralisation. Soit une population de n individus (i=1,…,n) partitionnée en k sous-populations (groupes d'âge, régions, niveau d'éducation, etc.). Chaque sous-population est de taille n_j (j=1,…,k). On note le revenu de l'individu i du groupe j : x_{ij}. Le vecteur de revenu global est défini par : $x = (x_{11}, x_{21}, ..., x_{ij}, ..., x_{nk})$.

L'agrégativité :

Une mesure agrégative est définie par :

$$I\ (x_{11}, x_{21}, ..., x_{ij}, ..., x_{nk}) = F^k\{I_1(x_{11}, x_{21}, ..., x_{n_1 1}), ..., I_k(x_{1k}, x_{2k}, ..., x_{n_k k});$$
$$Y_1, Y_2, ..., Y_k; n_1, ..., n_k\},$$

où I est l'indice d'inégalité associé à la population globale, μ_k la moyenne du groupe k, I_k l'indice d'inégalité associé au groupe k et Y_j la somme des revenus de la sous-population j :

$$Y_j = \sum_{i=1}^{n_j} x_{ij}\ .$$

L'agrégativité est une propriété générale indiquant qu'il n'est pas indispensable de connaître les caractéristiques exactes des distributions des sous-groupes pour déterminer le montant des disparités de revenu. Les mesures d'inégalité de chaque sous-population et les éléments agrégés tels que n_k et $n_k\mu_k$ sont simplement nécessaires. Les mesures d'inégalité agrégatives vérifient la décomposabilité élémentaire suivante :

$$\begin{aligned} I\ &= F^k\{I_1, ..., I_k; Y_1, Y_2, ..., Y_k; n_1, ..., n_k\} \\ &- F^k\{0, ..., 0; Y_1, Y_2, ..., Y_k; n_1, ..., n_k\} \\ &+ F^k\{0, ..., 0; Y_1, Y_2, ..., Y_k; n_1, ..., n_k\} \end{aligned}$$

où $F^k\{0,...,0; Y_1, Y_2, ..., Y_k; n_1, ..., n_k\}$ est un terme désignant une inégalité associée à des groupes égalitaires. Il est par conséquent possible de formuler les inégalités à l'intérieur des groupes et les inégalités entre les groupes. La contribution des inégalités intragroupes (I_w) à l'inégalité totale est définie comme la différence entre les inégalités observées dans la population mère et celles qui seraient effectivement observées si les individus du groupe j avaient tous le même revenu :

$$\begin{aligned} I_w = {} & F^k\{I_1,\ldots,I_k\,; Y_1, Y_2,\ldots, Y_k\,; n_1,\ldots, n_k\} \\ & - F^k\{0,\ldots, 0\,; Y_1, Y_2,\ldots, Y_k\,; n_1,\ldots, n_k\}. \end{aligned}$$

La contribution des inégalités intergroupes à l'inégalité totale est :

$$I_b = F^k\{0,\ldots, 0\,; Y_1, Y_2,\ldots, Y_k\,; n_1,\ldots, n_k\}.$$

La décomposition en sous-groupes de la mesure d'inégalité est donc :

$$I = I_w + I_b.$$

Pour définir les inégalités à l'intérieur d'un groupe particulier, on effectue la même opération que pour I_w. La contribution des inégalités associées au groupe j à l'inégalité totale s'écrit :

$$\begin{aligned} I_{wj} = {} & F^k\{I_1,\ldots, I_j,\ldots, I_k\,; Y_1, Y_2,\ldots, Y_k\,; n_1,\ldots, n_k\} \\ & - F^k\{I^{n_1},\ldots, 0,\ldots, I_k\,; Y_1, Y_2,\ldots, Y_k\,; n_1,\ldots, n_k\}. \end{aligned}$$

Cette définition peut paraître paradoxale, car la somme des contributions de chaque groupe I_{wj} n'est pas forcément égale à I_w. Les mesures ne sont donc pas indubitablement additives.

L'additivité :

Une mesure d'inégalité est additivement décomposable si elle vérifie la relation :

$$\sum_{j=1}^{k} I_{wj} = I_w.$$

Les mesures additivement décomposables sont donc agrégatives. L'additivité apparaît comme un cas particulier de l'agrégativité. Les propositions et définitions suivantes restreignent l'additivité.

Proposition 1.

Les mesures agrégatives dérivables s'expriment sous la forme suivante :

$$I = \sum_{j=1}^{k} f\left(Y_j, n_j\right) I_j + I\left(\mu_1,\ldots, \mu_1, \mu_2,\ldots, \mu_2,\ldots, \mu_k\right).$$

Cette proposition définit la fonction F^k comme une fonction séparable en chaque mesure I_j. A partir de cette proposition, il est possible de s'intéresser aux mesures homogènes de degré zéro, c'est à dire, aux indicateurs respectant l'invariance relative (confer les propriétés axiomatiques précédentes).

Proposition 2.

Les mesures décomposables, dérivables, homogènes de degré zéro et vérifiant le principe de population s'expriment sous la forme fonctionnelle suivante :

$$I = \sum_{j=1}^{k} g\left(z_j, v_j\right) I_j + I\left(\mu_1, \ldots, \mu_1, \mu_2, \ldots, \mu_2, \ldots, \mu_k\right),$$

où $v_j = n_j / n$, et z_j est la proportion de revenu total détenue par le groupe j : Y_j / Y.

Les mesures additivement décomposables sont donc des indicateurs d'inégalité qui s'écrivent sous la forme d'une moyenne pondérée des indices associés à chaque groupe (I_j) à laquelle on ajoute l'indicateur d'inégalité intergroupe. La proposition 2 ne prend pas en compte le principe de transfert de Pigou-Dalton. Il existe alors des répercutions sur l'indicateur de Gini. Si l'individu j transmet une partie de son revenu aux individus k et l, on doit avoir :

$$\frac{\Delta x_k}{\Delta x_l} = \frac{j-l}{k-j},$$

sinon l'indicateur de Gini n'est pas insensible à ce transfert. Les indices j, l et k renvoient au rang de l'individu dans la distribution lorsque celle-ci est classée par ordre décroissant. La restriction selon laquelle la valeur du transfert doit égaliser la différence de rang est due à la structure particulière de l'indice de Gini que l'auteur prend en compte :

$$G_{jj} = 1 + \frac{1}{n_j} - \left(\frac{2}{n_j^2 \mu}\right)\left(x_1, 2x_2, \ldots, n_j x_{n_j}\right),$$

où G_{jj} est le coefficient de Gini associé au groupe j.

Cette définition du coefficient de Gini ne permet pas de prendre en considération l'ensemble des transferts de Pigou-Dalton. On peut démontrer que certains transferts n'obéissent pas à la différence de rang entre les individus.[11] Il est alors important que les mesures décomposables respectent la dérivabilité, la symétrie (SM), l'invariance relative (IR), l'axiome de Pigou-Dalton et la configuration suivante :

[11] Remarquons que d'autres formules du coefficient de Gini prennent en compte l'ensemble des transferts (confer par exemple la revue de littérature sur l'indice de Gini ci-après).

$$I = \sum_{j=1}^{k} \frac{n_j}{n} I_j \left(x_{1j}, \ldots, x_{n_j j}\right) + I\left(\mu_1, \ldots, \mu_1, \mu_2, \ldots, \mu_2, \ldots, \mu_k\right). \qquad \text{(D)}$$

Cette forme fonctionnelle sera donc utilisée pour décrire les fonctions décomposables respectant l'ensemble des axiomes énoncés.

Proposition 3.

Une mesure dérivable, symétrique, et vérifiant la décomposabilité (*D*) s'écrit :

$$I = \frac{1}{n} \sum_{j=1}^{k} \left(K\left(x_j\right) - K\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right),$$

où K est une fonction dérivable.

Proposition 4.

Une mesure dérivable, symétrique, et vérifiant la décomposabilité (*D*) s'écrit :

$$L = log\left(\mu\right) - \frac{1}{n}\sum_{i=1}^{n} log\, x_i .$$

Cette mesure décomposable est séparable en pondérations de populations (conformément à l'équation (D) où le poids est la proportion de population du groupe j : $n_j\ /\ n$). Il existe des mesures où les pondérations font intervenir la part de revenu. Il s'agit des mesures décomposables pondérées par le revenu :

$$I = \sum_{j=1}^{k} \frac{Y_j}{Y} I_j \left(x_{1j}, \ldots, x_{nj\, j}\right) + I\left(\mu_1, \ldots, \mu_1, \mu_2, \ldots, \mu_2, \ldots, \mu_k\right). \qquad \text{(D1)}$$

Proposition 5.

La seule mesure d'inégalité dérivable, symétrique, vérifiant la décomposabilité (*D*1) *et les autres axiomes fondamentaux est l'indice de Theil*[12] *:*

$$T = \sum_{i=1}^{n} y_i \, log\, y_i \ \text{, où } \ y_i = x_i \Big/ \sum_{i=1}^{n} x_i .$$

Cette dernière proposition de Bourguignon (1979) caractérise le problème de décomposabilité en agrégativité et en décomposabilité additive. Les mesures d'inégalité sont séparées en indicateurs intragroupes et intergroupes. La somme des poids de la décomposition n'est pas égale à 1 pour les propositions 1, 2 et 3. Elle est égale à 1 pour

[12] Rappelons que l'indice de Theil est un cas particulier de l'entropie généralisée lorsque $c = 0$.

les propositions 4 et 5 [remarque de Deutsch et Silber (1999)]. L'agrégativité est moins restrictive car elle est définie sur les contributions des inégalités associées à chaque groupe (I_j). La décomposabilité additive est plus difficile à obtenir. Elle se base sur la mesure intragroupe (I_w). Plus précisément, l'élément intragroupe est défini par une moyenne pondérée des inégalités associées à chaque groupe (I_j). L'additivité necessite des propriétés axiomatiques plus restrictives, dont seul le coefficient de Theil respecte les conditions. Shorrocks (1980) va néanmoins démontrer que le coefficient de Theil n'est pas l'unique mesure.

I.3.2. La décomposabilité additive reconsidérée : Shorrocks (1980, 1984, 1988) :

L'hypothèse de décomposabilité additive s'exprime par :

$$I = \sum_{j=1}^{k} w_j(\mu, n) I_j(x_j) + I_b, \tag{DA}$$

où I_j représente l'indice d'inégalité associé au groupe j, w_j le poids attaché à l'indicateur I_j et I_b la mesure intergroupe. A partir de cette hypothèse, Shorrocks précise tout d'abord la proposition 3 de Bourguignon (1979).

Théorème 1. *Si I est continu, symétrique, normalisé*[13]*, dérivable* (*dérivées partielles de premier ordre continues*) *et vérifiant* (*DA*) *alors :*

$$I = \frac{1}{\theta(\mu, n)} \sum_{i=1}^{n} [\varphi(x_i) - \varphi(\mu)].$$

Ce théorème définit la classe entière des mesures additivement décomposables en imposant de sévères restrictions. L'intérêt de cette séparabilité est d'obtenir une généralisation des résultats de Bourguignon par l'entropie généralisée.[14]

Théorème 2. *L'indice d'inégalité I* (*possédant des dérivées secondes continues*) *satisfait l'axiome d'invariance* (*IR*) *et les axiomes* (*PD*), (*PP*), (*SM*), (*NM*) *et* (*DA*) *si I est défini par,*

[13] La normalisation concerne l'égalité des revenus. Si tous les membres de la société ont le même revenu, alors l'indice d'inégalité est égal à zéro.
[14] Voir aussi la caractérisation axiomatique de l'entropie et celle du concept d'additivité par Cowell et Kuga (1981).

$$I = \begin{cases} \sum_{i=1}^{n} \frac{A}{nc\,(c-1)} \left((x_i/\mu)^c - 1 \right), \forall c \neq 0,1 \\ A \sum_{i=1}^{n} \log \frac{\mu}{x_i} \\ A \sum_{i=1}^{n} \frac{x_i}{\mu} \log \frac{x_i}{\mu} \end{cases}$$

où A > 0.

On constate que ce théorème généralise les mesures additivement décomposables (et vérifie les propriétés énoncées) par un multiple de l'indice de l'entropie généralisée *Sc* :

$$Sc = \begin{cases} \sum_{i=1}^{n} \frac{\left((x_i/\mu)^c - 1 \right)}{nc\,(c-1)}, \forall c \neq 0,1 \\ \frac{1}{n} \sum_{i=1}^{n} \log \frac{\mu}{x_i}, \forall c = 0 \\ \frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{\mu} \log \frac{x_i}{\mu}, \forall c = 1. \end{cases}$$

Les poids de la décomposition de l'entropie généralisée sont donnés par :

$$w_j = \frac{n_j}{n} \left(\frac{\mu_j}{\mu} \right)^c.$$

La somme des poids est égale à l'unité uniquement pour $c = 0$ et $c = 1$. Par conséquent, l'indice intragroupe $\sum_j w_j(\mu, n) I_j(x_j)$ est une moyenne pondérée de l'indice I_j. Lorsque la somme des poids n'est pas égale à l'unité, l'expression $1 - \sum_j w_j$ est proportionnelle à l'inégalité intergroupe I_b. Cette proportionnalité indique que les poids de l'inégalité intragroupe ne sont pas indépendants de l'inégalité intergroupe [confer les travaux de Foster et Schneyerov (1997) sur l'indépendance]. Un autre problème survient lorsqu'on essaie d'interpréter la décomposition. On peut se demander quel serait le pourcentage de réduction de l'inégalité si les différences moyennes étaient éliminées. Mais si les moyennes sont modifiées, les poids intragroupes changent (car ils dépendent de la moyenne μ_j). Par conséquent, une dépendance existe aussi entre les inégalités intragroupes et les inégalités intergroupes. Pour éliminer cette dépendance, il est indispensable que les poids ne dépendent pas des moyennes. La seule mesure où les poids sont les proportions de population n_j/n (donc indépendants de la moyenne) correspond au cas particulier $c = 0$:

$$S_0 = \frac{1}{n} \sum_{i=1}^{n} \log \frac{\mu}{x_i}.$$

La contribution des inégalités intragroupes à l'inégalité totale S_0 est donc :

$$\sum_{j=1}^{k} w_j(\mu, n) I_j(x_j) = \sum_{j=1}^{k} \frac{n_j}{n} \sum_{i=1}^{n_j} \frac{1}{n_j} log \frac{\mu_j}{x_{ij}}.$$

La contribution des inégalités intergroupes à l'inégalité totale est donnée par :

$$I_b = \sum_{j=1}^{k} \frac{n_j}{n} log \frac{\mu_j}{\mu}.$$

Parallèlement, Cowell (1980a) élabore la même classe d'entropie généralisée (avec le paramètre β) :

$$I_\beta = \begin{cases} \dfrac{1}{2n} \sum_{i=1}^{n} \dfrac{x_i}{\mu}\left(\dfrac{x_i}{\mu} - 1\right), \forall \beta = 1 \\ \dfrac{1}{n} \sum_{i=1}^{n} \dfrac{x_i}{\mu} log \dfrac{x_i}{\mu}, \forall \beta = 0 \\ \dfrac{1}{n} \sum_{i=1}^{n} log \dfrac{\mu}{Mg}, \forall \beta = -1, \end{cases}$$

où Mg est la moyenne géométrique de la distribution. On retrouve les mêmes résultats que l'entropie de paramètre c. Le coefficient MLD équivaut au cas où $\beta = -1$, l'indice de Theil au cas où $\beta = 0$, et la moitié du carré du coefficient de variation au cas où $\beta = 1$. La réflexion de Cowell est cependant basée sur une hypothèse différente : la symétrie partielle. Il s'agit de l'axiome de symétrie (d'anonymat) qui s'exerce à l'intérieur de chaque sous-groupe. Les mesures d'inégalité satisfaisant la symétrie partielle, la dérivabilité au deuxième ordre, la décomposabilité (inégalité à l'intérieur des groupes et entre les groupes) et l'additivité sont dérivées de l'entropie généralisée. Russell (1985) démontre néanmoins que l'additivité et la décomposabilité sont suffisantes pour aboutir aux résultats de Cowell (1980a, 1980b).

Shorrocks (1984) propose une version plus faible de l'axiome (DA), en définissant le principe agrégatif :

$$I = f\left(I(x_j), \mu(x_j), n_j\right), \forall j = 1, ..., k \;;$$

où x_j est la distribution du groupe j et f une fonction continue en son premier argument. Il élabore par la suite le théorème 3 suivant qui ne prend plus en compte la dérivabilité.

Théorème 3. *I satisfait (IR), (PD), (PP), (SM), (NM) et le principe agrégatif si et seulement si (ssi)* $\exists\, h : [0, \infty[\to \Re_+,\ h(0) = 0 \mid h(I(x)) = Sc(x)$.

Le principe agrégatif prévoit la continuité de la fonction *f*. En 1988, Shorrocks va incorporer l'idée de cohérence. Le principe de décomposition cohérente va donc suppléer aux principes précédents.[15]

La décomposabilité cohérente en sous-groupes (monotonie). Une mesure d'inégalité vérifiant la décomposabilité cohérente s'écrit :

$$I = f\left(I_1,..., I_j,..., I_k\,; p_1,..., p_j,..., p_k\,; s_1,..., s_j,..., s_k\right),$$

où $p_j = \frac{n_j}{n}$, $s_j = \frac{n_j\ \mu_j}{n\ \mu}$ et f une fonction croissante avec ses *k* premiers arguments. Supposons que le groupe *j* voit ses revenus se modifier de telle sorte que la moyenne et le nombre d'individus du groupe *j* restent inchangés. La décomposition cohérente indique alors que si les inégalités au sein du groupe *j* augmentent, l'inégalité totale doit augmenter, *ceteris paribus* (les distributions des autres groupes restent les mêmes).

Théorème 4. *Un indice d'inégalité satisfait la continuité, les axiomes* (*PD*)*,* (*PP*)*,* (*SM*)*,* (*NM*)*,* (*IR*) *et la décomposition additive cohérente en sous-groupes si I est l'entropie généralisée Sc.*

L'ensemble des ces théorèmes montre que l'entropie généralisée rassemble d'importantes propriétés. Par contre, l'indice de Gini ne satisfait pas la décomposabilité additive. Même si l'indice peut s'écrire sous la forme d'une moyenne pondérée des indices associés aux groupes *j*, il est construit autour de trois éléments. Il viole aussi la propriété de décomposition cohérente en sous-groupes. En effet, les inégalités à l'intérieur d'un groupe peuvent augmenter en générant une baisse de l'inégalité totale. Cowell (1988) remarque à ce sujet que la mesure de Gini, la variance des logarithmes et l'écart moyen relatif[16] sont trois indicateurs d'inégalité qui ne respectent pas la

[15] Confer aussi Deutsch et Silber (1999) ou Cowell (2000) pour des explications claires sur la décomposition cohérente : « subgroup consistency ». Un concept similaire plus faible, celui de décomposition monotone, avait été introduit par Foster, Greer et Thorbecke (1984) dans le cadre des mesures de pauvreté. Il requiert simplement que le reste de la distribution ne bouge pas après une quelconque variation de la pauvreté dans au moins un groupe constituant la population globale.

[16] Rappelons que la variance des logarithmes s'écrit :

$\frac{1}{n}\sum_{i=1}^{n}\left(log\left(\frac{x_i}{Mg}\right)\right)^2$ où *Mg* est la moyenne géométrique. L'écart moyen relatif est :

$\frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i}{\mu}-1\right|$.

cohérence en sous-groupes. L'entropie généralisée satisfait donc d'importantes propriétés que les autres indicateurs n'intègrent pas. Ebert (1988) propose le théorème suivant pour doter l'indice de Gini de nouvelles propriétés.

Théorème 5 (Ebert, 1988). *I est défini sur l'ensemble des distributions classées par ordre non décroissant et satisfait l'invariance* (*IR*)*,* (*PD*)*,* (*PP*)*,* (*NM*)*, la différentiabilité et la décomposabilité additive pour des distributions qui ne se chevauchent pas si I est l'indice de Gini ou l'entropie généralisée Sc.*

L'indice de Gini est décomposable en trois éléments : un élément intragroupe, un élément intergroupe et une composante qui mesure l'intensité du chevauchement entre les distributions des sous-groupes. Lorsqu'il n'y pas de chevauchement entre les distributions, l'intensité du chevauchement est nulle, et l'indicateur de Gini satisfait la décomposabilité additive et la cohérence en sous-groupes.

Néanmoins, cette troisième composante apporte une information supplémentaire par rapport aux mesures de l'entropie. La revue de la littérature concernant l'indicateur de Gini démontre [notamment les travaux de Dagum (1997a, 1997b)] que l'indice de Gini possède une spécification économique et statistique intéressante.

II. Une revue(non exhaustive) des différentes décompositions de l'indice de Gini :

La décomposition de l'indice de Gini fut présentée pour la première fois en 1967 par Bhattacharya et Mahalanobis [même si pour certains, la première tentative revient à Soltow (1960)]. Trente ans plus tard, cette décomposition initiale est différente de la décomposition proposée par C. Dagum (1997a). Il n'existe pas de décomposition unique. Plusieurs méthodes sont recensées et basées sur des concepts mathématiques et statistiques différents. La somme des composantes de la décomposition est cependant toujours égale à l'indice de Gini global. Mais les éléments décomposés, obtenus par construction, ne possèdent pas les mêmes significations.

II.1. L'approche par la courbe de concentration de Bhattacharya et Mahalanobis (1967) :

En 1967, l'objectif de Bhattacharya et Mahalanobis est d'élaborer de nouveaux indicateurs statistiques pour mesurer les disparités régionales en Inde. Ils s'intéressent ainsi aux dépenses de consommation. La courbe de concentration[17] est au cœur de leur décomposition. Ils font référence, dans une note, à une décomposition en sous-groupes dont les distributions ne se chevauchent pas. Ils montrent que celle-ci n'est pas entièrement satisfaisante. En effet, la principale difficulté de la décomposition de l'indice de Gini est de tenir compte du chevauchement entre les distributions. Cette première approche soulève donc un problème fondamental, tout en montrant que la décomposition reste intéressante pour l'analyse des inégalités.

La courbe de concentration (la courbe de Lorenz) est construite pour la population indienne. Celle-ci est partagée en 12 classes. Les parts de consommation cumulatives de chaque classe sont données par :

$$Q_i = \sum_{j=1}^{i}\left(p_j \bar{x}_j \Big/ \sum_{j=1}^{12} p_j \bar{x}_j \right),\ \forall i = 1,\dots,12. \tag{1}$$

Le terme $\bar{x}_j$ représente la moyenne de consommation par personne (pour 30 jours). Le terme p_j est le pourcentage de population détenue par chaque classe. Les pourcentages de population cumulés correspondant aux Q_i sont :

$$P_i = \sum_{j=1}^{i} p_j,\ \forall i = 1,\dots,12. \tag{2}$$

La courbe de concentration s'obtient en liant successivement les points de coordonnées (P_i, Q_i). Le coefficient de concentration L (coefficient de Gini) est ensuite dérivé à partir de la somme des aires du trapèze[18] :

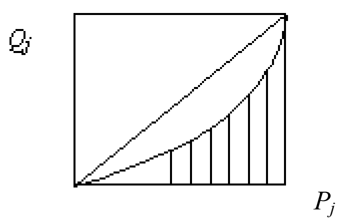
$$L = 1 - \sum_{j=1}^{12} p_j \left(Q_j + Q_{j-1}\right),\ \forall i = 1,\dots,12. \tag{3}$$

[17] La courbe de Lorenz est un cas particulier de la courbe de concentration. Voir la généralisation de Pyatt, Chen et Fei (1980).

[18] L'aire du trapèze est $T_j = \frac{\left(Q_j + Q_{j-1}\right)p_j}{2}$. La somme des aires des trapèzes est donc de :

$Aire = \sum_j \frac{\left(Q_j + Q_{j-1}\right)p_j}{2}$. L'indice de Gini ou de concentration est alors égal à :

$L = 1 - 2\sum_j \frac{\left(Q_j + Q_{j-1}\right)p_j}{2}$, où la hauteur du trapèze est p_j et non P_j.



Le coefficient de concentration équivaut à deux fois l'aire comprise entre la courbe de concentration et la première bissectrice. Pour effectuer la décomposition en sous-groupes ($\forall j = 1,\ldots, k$ groupes), la composante intergroupe est privilégiée.

Supposons une variable croissante non négative $(x_1,\ldots, x_k)$ associée aux probabilités $(p_1,\ldots, p_k)$ telles que $\sum_{j=1}^{k} p_j = 1$. La moyenne est donc : $\mu = \sum_{j=1}^{k} x_j p_j$. Par conséquent la fonction,

$$\Phi_r = \sum_{j=1}^{r} p_j x_j / \mu \,,\ \forall i = 1,\ldots,12\,, \qquad (4)$$

est la proportion de la variable x détenue par les individus pour lesquels $x \leq x_r$. En reliant Φ_r à la proportion d'individus $F_r = \Sigma_j p_j$, on obtient une courbe de concentration. La méthode repose sur une agrégation d'aires. Il manque donc une certaine précision par rapport au domaine continu. Une autre technique consisterait à utiliser la différence moyenne de Gini [Δ] divisée par 2μ. La différence moyenne de Gini est donnée par :

$$\Delta = E\,\left|x_1 - x_2\right|, \qquad (5)$$

où E est l'opérateur espérance et x_1 et x_2 représentent deux observations indépendantes de la population globale. L'expression Δ se réécrit :

$$\begin{aligned} \Delta &= 2 \sum_{x_j \leq x_i} p_i p_j \left|x_i - x_j\right| \\ &= 2 \sum_i p_i p_j \left(x_i p_i F_i - p_i \mu \Phi_i\right) \\ &= 2\mu \sum_i \left[F_i\left(\Phi_i - \Phi_{i-1}\right) - \Phi_i\left(F_i - F_{i-1}\right)\right]. \end{aligned} \qquad (6)$$

On retrouve la relation selon laquelle la valeur $\Delta/2\mu$ est égale à deux fois l'aire ente la courbe de Lorenz et la première bissectrice :

$$L = 1 - \sum_j p_j \left(\Phi_j - \Phi_{j-1}\right). \qquad (7)$$

Après avoir montré le lien entre la courbe de concentration et le coefficient de Gini (et la moyenne des différences de Gini Δ), le coefficient Δ est désagrégé en indicateurs intragroupes et intergroupes. L'indicateur intergroupe montre comment les inégalités globales de consommation peuvent être réduites si chaque distribution intragroupe est redistribuée de manière égalitaire. L'indice de concentration intergroupe est défini par $\Delta_B/2\mu$, où Δ_B est la moyenne des différences de Gini entre les groupes :

$$\Delta_B = \sum_{i \neq j} p_i p_j \left|\,\mu_i - \mu_j\right|, \qquad (8)$$

où μ_i et μ_j sont respectivement les moyennes des groupes i et j ($\forall i, j = 1,\ldots, k$) et k le nombre de sous-groupes. La moyenne des différences de Gini peut se récrire en fonction des k groupes :

$$\Delta = \sum_i p_i^2 \Delta_i + \sum_{i \neq j} p_i p_j \left\{ E \left| x_i^{(1)} - x_j^{(2)} \right| \right\}, \tag{9}$$

où $x_i^{(1)}$ et $x_j^{(2)}$ sont des distributions indépendantes, et Δ_i la moyenne des différences de Gini pour le groupe i. Etant donné les expressions Δ et Δ_B, il est possible de mesurer l'inégalité intragroupe :

$$\boxed{\begin{aligned} \Delta_W &= \Delta - \Delta_B \\ &= \sum_i p_i^2 \Delta_i + \sum_{i \neq j} p_i p_j \left\{ E \left| x_i^{(1)} - x_j^{(2)} \right| \right\} - \sum_{i \neq j} p_i p_j \left| \mu_i - \mu_j \right|. \end{aligned}} \tag{10}$$

Le deuxième terme de l'équation est nul si les distributions ne se chevauchent pas. Dans ce cas, la décomposition comporte deux éléments, au même titre que l'analyse de la variance. Il existe donc une courbe de concentration intergroupe et une courbe de concentration globale. L'aire située entre les deux courbes représente la mesure $\sum_i p_i^2 \Delta_i$. Il n'est cependant pas possible d'avoir une courbe de concentration pour l'élément intragroupe Δ_W. Par définition, il s'agit des variabilités à l'intérieur des groupes (comme la composante intragroupe de l'analyse de la variance). Mais cet élément n'a pas de spécification propre. La composante intragroupe est issue *ex post* de la détermination de la moyenne des différences de Gini globale et de la moyenne des différences de Gini intergroupe.

La première décomposition de l'indice de Gini soulève donc des difficultés. Les auteurs comparent la décomposition de l'indice de Gini à l'analyse de la variance qui comprend deux éléments. La décomposition offre trois composantes au lieu de deux. A partir de Bhattacharya et Mahalanobis la littérature va s'intéresser à cette décomposition particulière.

II.2. L'approche matricielle de Rao (1969) :

Rao utilise une approche matricielle pour montrer que la décomposition de l'indice de Gini s'écrit sous une forme quadratique. Cette expression permet d'identifier deux éléments. Le premier est une moyenne pondérée des indicateurs de concentration à l'intérieur des groupes. Le deuxième est basé sur les différences entre les sous-populations[19] en relation avec les revenus par tête. La décomposition est issue d'un découpage en classes de la population globale. Les notations sont les suivantes :

(i) p_{ij} est la proportion de population du groupe j ($\forall j = 1,\ldots, k$) appartenant à la classe i ($\forall i = 1,\ldots, t$) ;

(ii) x_{ij} est la proportion de revenu de la classe i de la sous-population j mesurée sur le revenu total da la population ;

(iii) $b_i = \sum_{i=1}^{t} x_{ij} \Big/ \sum_{i=1}^{t} p_{ij}$ est le rapport par tête de la $j^{\text{ème}}$ sous-population sur le revenu global.

Le coefficient de concentration associé à la population totale est donné par :

$$\begin{aligned} c &= \sum_{i=1}^{t-1}\left(\sum_{j=1}^{k} p_{ij} \times \sum_{j=1}^{k} x_{(i+1)\,j} - \sum_{j=1}^{k} p_{(i+1)\,j} \times \sum_{j=1}^{k} x_{ij}\right) \\ &= \sum_{i=1}^{t-1}\left(P_i\,. \times Q_{(i+1)}\,. - P_{(i+1)}\,. \times Q_i\,.\right), \end{aligned} \tag{11}$$

où $P_i\,. = \sum_i \left(\sum_{j=1}^{k} p_{ij}\right)$ et $Q_i\,. = \sum_i \left(\sum_{j=1}^{k} x_{ij}\right)$.

Vérifions la formule de Rao en considérant un découpage en deux classes. L'indice de Gini mesuré par la somme des aires des trapèzes de Bhattacharya et Mahalanobis est :

$$\begin{aligned} L &= 1 - \sum_{j=1}^{2} p_j \left(Q_j + Q_{j-1}\right) \\ &= 1 - p_1 Q_1 - p_2\left(Q_2 + Q_1\right). \end{aligned} \tag{12}$$

Etant donné que $p_1 = P_1$ et que $p_2 = P_2 - P_1$, alors :

$$\begin{aligned} L &= 1 - P_1 Q_1 - \left(P_2 - P_1\right)\left(Q_2 + Q_1\right) \\ &= 1 - P_2 Q_2 - P_2 Q_1 + P_1 Q_2 . \end{aligned} \tag{13}$$

Or, les valeurs cumulées P_2 et Q_2 sont égales à l'unité d'où :

$$L = -P_2 Q_1 + P_1 Q_2 . \tag{14}$$

D'après la formule de Rao, s'il existe deux classes ($i = 2$), alors :

$$c = \left(P_1 . \times Q_2\,. - P_2\,. \times Q_1 .\right). \tag{15}$$

On trouve bien l'égalité entre L et c.

[19] Le terme « sous-population », employé par l'auteur, désigne un sous-groupe de la population globale.

L'auteur montre qu'il est possible de mesurer les inégalités pour chacune des sous-populations j. Le coefficient devient :

$$c_{jj} = \sum_{i=1}^{t-1} \left(P_{ij}^* \times Q_{(i+1)j}^* - P_{(i+1)j}^* \times Q_{ij}^* \right), \tag{16}$$

où $P_{ij}^* = \sum_i \left(\sum_{j=1}^{k} p_{ij}^* \right)$, $Q_{ij}^* = \sum_i \left(\sum_{j=1}^{k} x_{ij}^* \right)$ et $p_{ij}^* = p_{ij} \Big/ \sum_{i=1}^{t} p_{ij}$, $x_{ij}^* = x_{ij} \Big/ \sum_{i=1}^{t} x_{ij}$. P^* et Q^* sont les proportions cumulées de taille et de revenu calculées à partir du groupe j. On retrouve par conséquent la cohérence avec l'équation (11). L'indice de concentration prend une nouvelle configuration si les groupes j sont différents. En effet, il est possible d'introduire le coefficient c_{jh} $(\forall j \neq h)$[20] tel que :

$$c_{jh} = \sum_{i=1}^{t-1} \left(P_{ij}^* \times Q_{(i+1)h}^* - P_{(i+1)j}^* \times Q_{ih}^* \right). \tag{17}$$

Les coefficients c_{jj} et c_{jh} peuvent être intégrés dans une matrice carrée C de taille $k \times k$. La diagonale de la matrice est composée des coefficients de Gini mesurés sur chaque groupe c_{jj}. Les éléments en dehors de la diagonale sont les indices c_{jh}. Soit p le vecteur (de taille k) des proportions de taille des différents groupes. L'élément p_j du vecteur p est donc la taille relative du groupe j par rapport à la population totale :

$$p_j = \sum_{i=1}^{t} p_{ij}. \tag{18}$$

Soit q le vecteur (de taille k) représentant les proportions de revenu de chaque groupe par rapport à la population mère. L'élément q_j du vecteur q s'écrit donc :

$$q_j = \sum_{i=1}^{t} q_{ij}. \tag{19}$$

L'indicateur de concentration, mesuré sur la population globale, peut donc se réécrire sous la forme matricielle suivante :

$$c = p' C\, q. \tag{20}$$

L'indice peut être défini à l'aide de deux autres matrices de même taille $k \times k$. Soit la matrice W composée des indices de concentration à l'intérieur des groupes. Les éléments de cette matrice sont définis par :

$$w_{jh} = c_{jj}. \tag{21}$$

La première ligne de W est composée de l'indicateur intragroupe c_{11} (et ainsi de suite). La deuxième matrice de taille $k \times k$ est notée D. Les éléments de D sont définis tels que :

[20] L'auteur ne donne pas de définition du terme c_{jh}, mais intuitivement on comprend qu'il s'agit du coefficient de Gini entre les groupes j et h. Cette idée sera notamment reprise par Dagum (1987, 1997a) dont la décomposition est issue de la spécification des indicateurs de Gini intragroupes et intergroupes.

$$d_{jh} = \begin{cases} 0 \ \forall j = h \\ c_{jh} - c_{jj} \ \forall j \neq h. \end{cases} \tag{22}$$

La relation qui unit les trois matrices est donc :

$$C = W + D.$$

L'indice de Gini se reformule par :

$$c = p' W q + p' D q. \tag{23}$$

Soit le vecteur $\underline{b}$ composé des éléments $b_i = \sum_{i=1}^{t} x_{ij} \Big/ \sum_{i=1}^{t} p_{ij}$, et $\underline{c}$ le vecteur des coefficients de concentration à l'intérieur des groupes (c_{jj}). Il s'ensuit la décomposition suivante :

$$c = p'\underline{c} + p'(\underline{b}D). \tag{24}$$

L'équation (24) montre que le coefficient de Gini est décomposé en deux termes. Le premier terme est une moyenne pondérée des indices de concentration à l'intérieur des groupes. Le second terme est une forme quadratique en p' où ($\underline{b}D$) est la matrice de la forme quadratique. Chaque paire distincte de sous-populations contribue deux fois à la forme quadratique. Cette dernière est dépendante d'un paramètre d. Ce paramètre mesure les différences entre les classes des distributions j et h. L'appellation « composante intergroupe » n'est donc pas acceptable.[21] En effet, cette expression est réservée aux indices qui reflètent les différences moyennes entre les groupes, alors que cette décomposition procure des inégalités entre les paires de groupes.

En définitive, la décomposition du coefficient de Gini réalisée par Rao dépend donc des inégalités à l'intérieur des groupes et des différences entre chaque paire distincte de sous-populations.

II.3. L'approche matricielle et la théorie des jeux de Pyatt (1976) :

L'article de Pyatt (1976) marque une rupture dans la littérature. L'expression « coefficient de concentration » est abandonnée pour l'expression « coefficient de Gini ». Le coefficient de Gini peut être interprété comme une valeur espérée, au sens de l'espérance mathématique. Cette espérance est basée sur le fait que chaque individu peut comparer son gain avec celui d'un autre individu qui serait tiré au hasard dans la population. Il est alors possible de fournir une nouvelle décomposition de l'indicateur de Gini basée sur la valeur conditionnelle espérée du jeu. Cette approche est liée à celle

[21] La composante intergroupe est basée sur les différences moyennes entre les groupes. Cette décomposition offre des différences entre paires de populations concernant les différentes classes.

de Bhattacharya et Mahalanobis. Elle est, en effet, construite sur les différences absolues de revenu. Soient n revenus notés : $x_1,\ldots, x_n$. L'indicateur de Gini faisant intervenir les comparaisons interpersonnelles entre paires d'individus est :

$$G = \frac{\left(1/2n^2\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{\left(1/n\right)\sum_{i=1}^{n} x_i}. \tag{25}$$

En développant l'expression suivante,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right| = 2\sum_{i=1}^{n}\sum_{j=1}^{n}\max\left(0,\ x_i - x_j\right), \tag{26}$$

l'indice de Gini se réécrit :

$$G = \frac{\left(1/2n^2\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\max\left(0,\ x_i - x_j\right)}{\left(1/n\right)\sum_{i=1}^{n} x_i}. \tag{27}$$

Cette dernière équation constitue la base à partir de laquelle le jeu peut être défini. Chaque individu est convié à cette expérience. Un revenu x est tiré au hasard dans la population totale. Si le revenu sélectionné x est supérieur à celui du joueur, alors ce dernier peut choisir x, sinon il peut garder son revenu actuel. Aucun individu, qui participe à ce genre de jeu, ne peut perdre. Toutes les personnes ont une probabilité non négative d'y gagner, excepté les gens très riches dont la probabilité se rapproche de zéro. Pour chaque individu, soit le gain est nul (s'il est riche) soit il est égal à la différence entre x et son revenu. Le gain espéré de l'individu i est donc :

$$\frac{1}{n}\sum_{j=1}^{n}\max\left(0,\ x_j - x_i\right) \geq 0. \tag{28}$$

En mesurant la moyenne des gains espérés individuels (m), on obtient le numérateur de l'indicateur de Gini :

$$m = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max\left(0,\ x_j - x_i\right) \geq 0. \tag{29}$$

L'indice de Gini est donc interprété comme la moyenne du gain espéré, si chaque individu a le choix entre être lui-même ou prendre la place d'une autre personne prise au hasard dans la société.

Supposons maintenant que la population soit décomposée en k sous-groupes. La moyenne du gain espéré se formule par :

$$m = \sum_{j=1}^{k}\sum_{i=1}^{k} E(\text{gain}/\, i \to j) Pr\,(i \to j), \tag{30}$$

où l'événement $i \to j$ désigne le fait qu'un individu du groupe i se compare aux membres du groupe j tirés au hasard. Par conséquent, l'expression E(gain / $i \to j$) représente la moyenne des gains espérés des individus appartenant au groupe i qui se comparent aux membres de j. Comme les comparaisons sont aléatoires, le terme $\text{Pr}\ (i \to j)$ peut être spécifié par :

$$\text{Pr}\ (i \to j) = \ p_i\, p_j \text{, tel que } \sum_{i=1}^{n} p_i = 1 \ , \forall i, j = 1, \ldots, k\ . \tag{31}$$

Le terme p_i est la proportion de la population globale qui se trouve dans le groupe i. On note $\underline{p}$ le vecteur colonne (de taille k) des proportions de population et $\underline{\mu}$ le vecteur colonne (de taille k) où le $k^{\text{ième}}$ élément est la moyenne du groupe k. Soit E la matrice ($k\times k$) où $e_{ij} = E$(gain / $i \to j$). L'indicateur de Gini peut donc se réécrire sous la forme matricielle suivante :

$$G = (\mu' p)^{-1} p' \text{E}\, p\ . \tag{32}$$

On remarque que $(\mu' p)^{-1}$ est le dénominateur de l'indice de Gini [équation (27)] : la moyenne des revenus de la population mère. Le deuxième terme $p'\text{E}\,p$ représente le numérateur de G (le terme m) : la moyenne des gains espérés individuels.

Soit E* : la matrice E normalisée. Le facteur de normalisation est la matrice $\hat{\mu}^{-1}$. Celle-ci est composée des vecteurs μ. Chaque élément de E est donc divisé par la moyenne des revenus de la population correspondante :

$$\text{E*} = \hat{\mu}^{-1}\,\text{E}\ . \tag{33}$$

Par conséquent, les éléments de la matrice E* se trouvant sur la diagonale principale sont les coefficients de Gini des k sous-populations. L'indicateur de Gini mesuré sur la population globale peut alors se réécrire :

$$G = \pi'\ \text{E*}\ p\ ,[22] \tag{34}$$

où $\pi = (m'p)^{-1}\ \hat{\mu}\ p$ est le vecteur des proportions de revenu agrégés de chaque groupe i.

Supposons maintenant qu'il n'y ait pas d'inégalités intragroupes. La matrice E* est uniquement composée de zéro sur la diagonale principale. Comme il n'est pas intéressant d'obtenir le revenu d'un groupe plus pauvre, les éléments sous la diagonale de E* sont nuls. Les éléments au-dessus de la diagonale sont : $\mu_j - \mu_i > 0$, car le gain

[22] Cette forme est équivalente à celle introduite par Rao (1969) : $c = p' C\, q$, équation (20).

associé à la chance de rejoindre un groupe plus riche équivaut à la différence des moyennes entre les groupes. De ce fait, E se réécrit :

$$\mathrm{E} = A'\hat{\mu} - \hat{\mu}A', \tag{35}$$

où la matrice A est une matrice ($k \times k$) dont les éléments au-dessus de la diagonale sont nuls et les autres égaux à un. Il s'agit simplement d'un cas particulier où il n'existerait pas d'inégalités intragroupes. Les inégalités intergroupes sont donc définies par les différences entre les moyennes des sous-groupes. Pour faire apparaître ces différences de moyenne, la matrice E issue de l'équation (32) devient :

$$\mathrm{E} = \mathrm{E}_1 + A'\hat{\mu} - \hat{\mu}A', \tag{36}$$

où E_1 est une matrice symétrique ($k \times k$) dont les éléments $e_{1ij} = \min\{e_{ij}, e_{ji}\}$. La matrice des gains espérés est donc séparée en deux composantes dont l'une ($A'\hat{\mu} - \hat{\mu}A'$) est basée sur les différences moyennes entre les groupes. La matrice E* se réécrit :

$$\mathrm{E}^* = \hat{\mu}^{-1}\mathrm{E}1 + \hat{\mu}^{-1}A'\hat{\mu} - A'. \tag{37}$$

Par conséquent, G devient :

$$\boxed{\begin{aligned} G &= \pi'\mathrm{E}^* p \\ &= \pi'\left[\hat{\mu}^{-1}\mathrm{E}1 + \hat{\mu}^{-1}A'\hat{\mu} - A'\right]p \\ &= \pi'\left[\mathrm{E}_1^* + \hat{\mu}^{-1}A'\hat{\mu} - A'\right]\mathrm{p}, \end{aligned}} \tag{38}$$

où $\mathrm{E}_1^* = \hat{\mu}^{-1}\mathrm{E}_1$.

L'équation (38) correspond à la désagrégation finale du coefficient de Gini. Cette décomposition comporte deux éléments. Les éléments sur la diagonale de E_1^* représentent les indicateurs de Gini associés à chaque sous-groupe. Tous les éléments hors diagonale sont égaux à zéro lorsque les distributions ne se chevauchent pas. Le deuxième élément concerne les inégalités entre les groupes : les différences entre les moyennes des distributions.

Cette décomposition est proche de celle de Bhattacharya et Mahalanobis (1967) dans le sens où les inégalités de chevauchement sont incluses dans la partie des inégalités intragroupes. Elle est légitime dans la mesure où les inégalités de chevauchement (les différences de revenu entre les groupes issues du chevauchement entre les distributions) sont dépendantes des inégalités intragroupes. En effet, plus le chevauchement augmente et plus les inégalités intragroupes sont faibles. Néanmoins, il est possible de séparer, comme l'a fait Rao (1969), l'indice de Gini en deux composantes, sans faire intervenir les inégalités intergroupes (les différences moyennes entre les groupes). Rappelons que Rao a décomposé l'indicateur en un élément intragroupe et un élément qui explique les différences de revenu entre les différentes classes des sous-groupes. Das et Parikh (1982) montrent que Bhattacharya-Mahalanobis et Pyatt spécifient d'abord les inégalités intergroupes et dérivent ensuite l'inégalité intragroupe comme un résidu, alors que Rao spécifie la composante intragroupe et dérive l'élément intergroupe comme un résidu [néanmoins défini comme les différences entre les classes (c_{ij})] au même titre que les décompositions de Mangahas (1975) et Soltow (1960).

II.4. L'approche de Mookherjee et Shorrocks (1982) et le rejet de la troisième composante issue du chevauchement entre les distributions :

On considère une population mère composée de n individus (ou ménages). On note μ la moyenne des revenus de cette population. Soit x_i le revenu du $i^{\text{ème}}$ individu ($\forall i = 1,\ldots, n$). On suppose que la population est divisée en k sous-populations de taille n_h et de moyenne μ_h ($\forall h = 1,\ldots, k$). N_h désigne l'ensemble des revenus du groupe h. L'indicateur de Gini s'écrit :

$$G = \frac{1}{2\mu n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right| . \tag{39}$$

Les différences de revenu peuvent être regroupées afin de mettre en exergue celles entre individus appartenant aux mêmes groupes et entre individus appartenant à des groupes différents :

$$\begin{aligned} G &= \frac{1}{2\mu n^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j=1}^{n_h}\left|x_{rh} - x_{jh}\right| + \sum_{r=1}^{n_h}\sum_{j\notin N_h}\left|x_{rh} - x_j\right|\right) \\ &= \frac{1}{2\mu n^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j=1}^{n_h}\left|x_{rh} - x_{jh}\right|\right) + \frac{1}{2\mu n^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j\notin N_h}\left|x_{rh} - x_j\right|\right), \forall h = 1,\ldots,k \ . \end{aligned} \tag{40}$$

On constate que l'indice est décomposé en deux éléments. Le premier représente les inégalités issues des différences entre les revenus des mêmes groupes. Il incorpore les différences binaires issues de chaque groupe. Elles représentent le numérateur des indicateurs de Gini de chaque groupe. Le deuxième est constitué des différences de revenu entre individus appartenant à des groupes différents. Pour mettre en évidence ces indices, il est nécessaire de faire apparaître le dénominateur : $1/2\mu_h n_h^2$. En multipliant et en divisant par ce terme on obtient :

$$G=\frac{1}{2\mu n^2}\times 2\mu_h n_h^2\times\frac{1}{2\mu_h n_h^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j=1}^{n_h}\left|x_{rh}-x_{jh}\right|\right)+\frac{1}{2\mu n^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j\notin N_h}\left|x_{rh}-x_j\right|\right),\ \forall h=1,\ldots,k\ ,\tag{41}$$

avec,

$$G_h=\frac{1}{2\mu_h n_h^2}\left(\sum_{r=1}^{n_h}\sum_{j=1}^{n_h}\left|x_{rh}-x_{jh}\right|\right),\tag{42}$$

représentant l'indice de Gini associé au groupe *h*. La décomposition de l'indice de Gini s'écrit alors :

$$G=\sum_{h=1}^{k}\left(\frac{n_h}{n}\right)^2\frac{\mu_h}{\mu}G_h+\frac{1}{2\mu n^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j\notin N_h}\left|x_{rh}-x_j\right|\right),\ \forall h=1,\ldots,k\,.\tag{43}$$

Supposons que les distributions *z* et *h* ne se chevauchent pas. La somme des différences binaires entre les revenus des individus de *z* et de *h* est :

$$\sum_{r=1}^{n_z}\sum_{j=1}^{n_h}\left|x_{rz}-x_{jh}\right|=n_z n_h\left|\mu_z-\mu_h\right|.\tag{44}$$

La décomposition de l'indice de Gini pour des distributions qui ne se chevauchent pas est donnée par :

$$G=\sum_{h=1}^{k}\left(\frac{n_h}{n}\right)^2\frac{\mu_h}{\mu}G_h+\frac{1}{2}\sum_z\sum_h\frac{n_z}{n}\frac{n_h}{n}\left|\frac{\mu_z}{\mu}-\frac{\mu_h}{\mu}\right|,\ \forall h\neq z=1,\ldots,k\tag{45}$$

Si la proportion d'individus appartenant au groupe *h* est notée $v_h = n_h / n$, et si la proportion de revenu détenue par le groupe *h* est $\lambda_h = \mu_h/\mu$, alors la décomposition de l'indice de Gini pour des distributions qui ne se chevauchent pas devient :

$$G=\sum_{h=1}^{k}(v_h)^2\lambda_h G_h+\frac{1}{2}\sum_z\sum_h v_z v_h\left|\lambda_z-\lambda_h\right|,\ \forall h\neq z=1,\ldots,k\,.\tag{46}$$

A contrario, la décomposition du coefficient de Gini pour des distributions qui se chevauchent peut s'écrire :

$$\boxed{G=\sum_{h=1}^{k}(v_h)^2\lambda_h G_h+\frac{1}{2}\sum_z\sum_h v_z v_h\left|\lambda_z-\lambda_h\right|+R,\ \forall h\neq z=1,\ldots,k},\tag{47}$$

où *R* dépend de l'intensité avec laquelle les distributions se chevauchent et est appelé « terme d'interaction ». Les auteurs ne spécifient pas le terme d'interaction. Il s'obtient simplement par différence :

$$R=\frac{1}{2\mu n^2}\sum_{h=1}^{k}\left(\sum_{r=1}^{n_h}\sum_{j\notin N_h}^{n_h}\left|x_{rh}-x_j\right|\right)-\frac{1}{2}\sum_z\sum_h v_z v_h\left|\lambda_z-\lambda_h\right|,\ \forall h\neq z=1,\ldots,k\,.\tag{48}$$

Lorsque les revenus sont également distribués à l'intérieur de chaque groupe, l'indicateur de Gini s'exprime sous la forme suivante :

$$G = \frac{1}{2}\sum_{z}\sum_{h} v_z v_h \left|\lambda_z - \lambda_h\right|, \ \forall h \neq z = 1,..., k . \quad (49)$$

Ce cas particulier élimine les inégalités à l'intérieur des groupes. Pour l'illustrer, les auteurs prennent l'exemple des tranches d'âge : il n'existe pas d'inégalités entre les individus d'une même classe d'âge. Une parfaite égalité des revenus prévaut dans chaque groupe. Les chevauchements entre les distributions sont donc inexistants (R = 0). L'inégalité totale est ainsi expliquée par les inégalités intergroupes : les différences entre les revenus moyens des classes d'âge. En revanche, la décomposition du coefficient de Gini (47) reste particulière. Le terme R est impossible à interpréter avec précision. Il s'agit simplement du terme qui permet de maintenir l'exactitude de la décomposition. De plus, celle-ci répond difficilement aux changements temporels. Si on considère une augmentation des inégalités intragroupes sans changement du nombre d'individus et de la moyenne de chaque classe, on s'attend donc à une augmentation de l'indice G. Le terme d'interaction peut néanmoins absorber cette variation pour faire diminuer G .[23]

Selon les auteurs, cette décomposition s'inscrit dans la lignée des travaux de Bhattacharya et Mahalanobis, Rao et Pyatt. Il s'agit en effet d'un mélange de ces décompositions. Rappelons que les décompositions de Bhattacharya et Mahalanobis (1967) et de Pyatt (1976) sont identiques dans le sens où elles spécifient l'indice intergroupe, alors que celle de Rao spécifie l'indice intragroupe. On remarque que Mookherjee et Shorrocks parviennent à préciser les fondements intragroupes et intergroupes de la décomposition. Par conséquent, il s'agit d'une synthèse des décompositions précédentes. Cette approche rejette cependant la décomposition de l'indicateur de Gini en trois éléments à cause de la non-spécification du terme d'interaction. Ce dernier trouve cependant une définition intéressante dans l'approche de Lerman et Yitzhaki (1991).

<u>II.5. L'approche de Lerman et Yitzhaki (1991) : une spécification de la troisième composante comme indice de stratification :</u>

[23] Il s'agit d'un exemple du non-respect de la décomposition cohérente en sous-groupes.

Les sociologues se consacrent davantage à l'étude de la stratification qu'à celle des inégalités de revenu. Il existe différentes définitions de la stratification. La stratification sociale est le partage de la société en groupes qui possèdent des éléments culturels et historiques différents. Elle fait appel à la notion de strates ou de couches horizontales regroupant une ou plusieurs caractéristiques de la population globale. Selon cette dernière définition, un indice mesurant la stratification doit capter le degré de chevauchement entre plusieurs distributions. Il existe donc un lien étroit entre les inégalités intergroupes et la notion de stratification entre les groupes. Ainsi, l'augmentation des inégalités d'un groupe pauvre (par augmentation des hauts revenus) peut augmenter la stratification. Une augmentation des inégalités des populations riches (par diminution des bas revenus) peut aussi accroître la stratification. Les décompositions qui lient les éléments intragroupes et intergroupes possèdent donc des éléments pouvant expliquer ce phénomène. Les mesures décomposables issues de l'entropie ne sont pas adéquates pour cette approche. En effet, les mesures intergroupes basées sur les différences entre les moyennes des groupes ne fournissent pas ce genre d'explication. A contrario, l'indice de Gini représente bien cette analyse.

Soient x_{ij} le revenu de l'individu i appartenant au groupe j., n le nombre d'individus total et n_j le nombre d'individus appartenant au groupe j tel que :

$$n = \sum_{j=1}^{k} n_j .$$

Soient p_j la proportion de la population appartenant au groupe j : $p_j = n_j/n$ et $F_Q(x_{ij})$ la fonction de répartition de la population globale Q. Elle mesure le rang du $i^{\text{ème}}$ individu (du groupe j) dans la population globale : $F_Q(x_{ij}) = \text{rang}\ (x_{ij})/n$.

Soit $F_j(x_{ij})$ la fonction de répartition du groupe mesurant le rang de l'individu i (du groupe j) dans la sous-population j : $F_j(x_{ij}) = \text{rang}\ (x_{ij})/n_j$. L'indice de stratification concerne différentes strates de la population et donne une mesure pour chaque groupe. La population doit être divisée en deux groupes : les individus issus du groupe j et les individus non issus du groupe j : « non j ». La fonction de répartition des « non j » ($-j$ en abrégé) est définie par : $F_{-j}(x_{ij}) = \text{rang}\ (x_{ij})/(n - n_j)$. Cette fonction classe tous les individus de la population globale exceptés ceux de la population j. La fonction de répartition de la population globale peut donc se réécrire comme une moyenne pondérée du classement des individus du groupe j et du reste de la population : $F_Q(x_{ij}) = p_j\, F_j(x_{ij}) + (1 - p_j)F_{-j}(x_{ij})$. On note $\text{Cov}_j(x,y)$ la covariance entre les variables x et y des membres du groupe j. L'indice de stratification S peut alors se formuler :

$$S_j = \frac{\mathrm{Cov}\left(\left(F_j - F_{-j}\right), x\right)}{\mathrm{Cov}\left(F_j, x\right)}. \tag{50}$$

Le numérateur exprime la covariance entre la variable revenu et la différence entre le rang de l'individu dans son propre groupe et celui qu'il aurait eu dans le reste de la population. Le dénominateur est un facteur de normalisation. L'indicateur de stratification est donc compris dans l'intervalle [-1,1].

- Si S_j = 1, aucun membre des autres groupes ne possède un revenu dont le montant se situe dans l'intervalle de revenu du groupe *j*. Seul, le groupe *j* occupe une place particulière dans la distribution. Il forme une strate parfaite.
- S_j diminue lorsque le nombre d'individus des autres groupes qui tombent dans l'intervalle de revenu du groupe *j* augmente. L'indicateur est donc sensible au chevauchement entre les distributions.
- Si S_j = 0, le rang de l'individu associé à son propre groupe est égal au rang mesuré sur le reste de la population ou sur la population globale. Dans ce cas, le groupe *j* ne forme pas de strates.
- Si S_j < 0, la divergence de classement des individus du groupe *j* dans la population globale est plus importante que celle qui prévaut dans le groupe *j*.
- Si S_j = -1, et si le groupe *j* est composé de deux groupes, cela signifie que les membres des deux groupes sont identiques et qu'ils sont situés aux extrémités de la distribution globale. Le groupe *j* est donc formé de deux strates parfaites.

Une société est donc stratifiée si l'indicateur est supérieur à zéro : il reflète la séparation entre les membres du groupe *j* et ceux du reste de la population. Pour relier la notion de stratification et la notion de chevauchement entre les groupes on construit l'indicateur de chevauchement O :

$$O_j = \frac{1-\left(1-p_j\right)S_j}{p_j}. \tag{51}$$

Cet indicateur ne prend pas en compte la part d'individus associée au groupe *j*. Plus la taille du groupe est petite et plus le coefficient O_j est important. En introduisant l'expression de l'indice de stratification dans l'équation, l'indicateur de chevauchement devient :

$$O_j = \frac{\mathrm{Cov}_i\left(x, R\right)}{\mathrm{Cov}_i\left(x, r_j\right)}, \tag{52}$$

où R est le rang des individus dans la population globale et r_j le rang dans la population j.[24] On obtient :

- $O_j = 1$ si $S_j = 1$;
- $O_j = 1/p_j$ si $S_j = 0$;
- $O_j = (2-p_j) / p_j$ si $S_j = -1$.

La stratification et le chevauchement sont donc liés, il est possible d'établir le lien avec la décomposition de l'indicateur de Gini. Ainsi, Lerman et Yitzhaki (1984) ont introduit la formule du coefficient de Gini basée sur l'opérateur de covariance :

$$G = \frac{2\text{Cov}\left(x, F(x)\right)}{\mu}, \tag{53}$$

où μ est la moyenne des revenus de la population globale.

Soit G_j l'indice de Gini mesuré sur le groupe j, $\bar{r}_j = \sum_{i=1}^{n_j} r_{ij} \Big/ n_j$ le rang moyen des individus du groupe j dans la population globale, et $F_j = \bar{r}_j / n$. La décomposition de l'indice de Gini devient :

$$\boxed{G = \sum_{j=1}^{k} p_j \frac{n_j \mu_j}{n\mu} G_j + \sum_{j=1}^{k} \frac{2\,\text{Cov}\left(\mu_j, F_j\right)}{\mu} + \sum_{j=1}^{k} p_j \frac{\mu_j}{\mu} S_j \left(p_j - 1\right).} \tag{54}$$

La première composante représente l'inégalité intragroupe : la moyenne pondérée des indices de Gini à l'intérieur des groupes. La deuxième composante est l'inégalité intergroupe. La troisième synthétise l'impact de la stratification. L'indicateur de Gini global équivaut à deux fois la covariance entre le revenu de l'individu et son rang dans la population globale, divisé par la moyenne des revenus (cf 53). Si les groupes deviennent des observations, l'indicateur intergroupe est défini de manière similaire. Il s'agit de deux fois la covariance entre la moyenne des revenus de chaque groupe et la moyenne des rangs, divisé par la moyenne totale des revenus. Pourtant, la lecture de la littérature montre que cet indice de Gini intergroupe est totalement différent de ceux introduits dans les approches précédentes. Ces dernières sont basées sur la moyenne de chaque sous-population. Si tous les groupes ont la même moyenne, l'indice intergroupe est nul. Dans la présente décomposition, la mesure d'inégalité intergroupe de Gini est nulle si et seulement si les groupes ont le même rang moyen. Il est donc possible d'avoir une composante intergroupe nulle, alors que les moyennes sont différentes. Dans le

[24] Confer l'approche de Milanovic et Yitzhaki (2002) pour une application internationale.

même contexte, l'indice intergroupe de Gini défini par Pyatt (1976) est positif. Dans l'approche de Lerman et Yitzhaki, les groupes sont égaux en terme de position moyenne à l'intérieur de la distribution. De ce fait, l'indice de Gini intergroupe peut revêtir une valeur négative. Par exemple, supposons que la population soit constituée de deux groupes. Le groupe pauvre comprend uniquement des faibles revenus et un seul individu extrêmement riche. Le rang moyen du groupe pauvre (dans la population totale) peut être inférieur à celui du groupe riche alors que la moyenne est plus élevée.

L'élément de stratification permet donc de proposer une décomposition dotée d'une troisième composante totalement définie. D'ailleurs, les auteurs critiquent l'approche de Mookherjee et Shorrocks (1982) selon laquelle le dernier élément est un terme d'interaction dépourvu de toute signification économique. La troisième composante est ici clairement définie. Une forte stratification implique une faible variabilité de rang. Celle-ci engendre à son tour un faible niveau d'inégalité. Une augmentation de la stratification exerce un effet négatif sur les inégalités. Il existe donc une relation opposée entre ces deux éléments.

II.6. L'approche par les courbes de Lorenz de Lambert et Aronson (1993) :

S'inspirant des travaux de Mookherjee et Shorrocks (1982), Lambert et Aronson proposent une nouvelle décomposition de la mesure de Gini. Ils s'intéressent surtout à ce que Mookherjee et Shorrocks nomment « terme d'interaction », ou « terme résiduel ». Rappelons que leur décomposition est la suivante :

$$\begin{aligned} G &= \sum_{h=1}^{k}\left(\frac{n_h}{n}\right)^2 \frac{\mu_h}{\mu}\, G_h + \frac{1}{2}\sum_z\sum_h \frac{n_z}{n}\frac{n_h}{n}\left|\frac{\mu_z}{\mu} - \frac{\mu_h}{\mu}\right| + R,\ \forall h \neq z = 1,\ldots,k \\ &= \sum_{h=1}^{k} a_h G_h + G_B + R,\ \forall a_h = \left(\frac{n_h}{n}\right)^2 \frac{\mu_h}{\mu} \qquad (55) \\ &= G_W + G_B + R. \end{aligned}$$

La première composante (G_W) est l'inégalité intragroupe, la deuxième (G_B) est l'inégalité intergroupe et la troisième (R) est le terme résiduel (qui est égal à zéro si les distributions ne se chevauchent pas). En partant d'une simple approche géométrique basée sur la courbe de Lorenz, les auteurs spécifient les trois composantes de la décomposition.

La population mère est composée de n individus dont les revenus sont notés : x_1, $x_2,\ldots, x_n$. Cette population comprend k groupes. La moyenne de la population est μ. La moyenne et la taille de chaque groupe sont respectivement μ_j et n_j ($\forall j = 1,\ldots, k$). La courbe de Lorenz est définie par $L(P)$. L'ordonnée $L(P)$ représente la part des revenus cumulés de la population globale détenue par chaque classe. L'abscisse P est la proportion de population totale de chaque classe. Le coefficient de Gini se formule :

$$G = 2\int_0^1 \left(P - L(P)\right) dP \,. \tag{56}$$

Les termes $P - L(P)$ définissent les segments situés entre la première bissectrice et la courbe de Lorenz. Si on multiplie par deux la somme des segments entre 0 et 1 , on obtient bien le coefficient de Gini : deux fois l'aire contenue entre la courbe de Lorenz et la bissectrice. A partir de la courbe de Lorenz, les étapes de la décomposition de l'indicateur de Gini sont les suivantes.

1. *La contribution des inégalités intergroupes*. On considère que les revenus sont également distribués à l'intérieur de chaque groupe. Les individus possèdent donc la moyenne de leur sous-groupe correspondant. Celles-ci sont classées par ordre croissant : le groupe 1 possède la plus faible moyenne, et ainsi de suite : $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$. La courbe de Lorenz intergroupe L_B est construite en imaginant que les individus du groupe j possèdent un revenu égal à μ_j, pour tous les $j = 1,\ldots, k$.

2. *La contribution des inégalités intragroupes*. Les revenus des individus sont classés par ordre croissant. On relie la proportion de population totale des individus du groupe j avec leur proportion de revenu pour obtenir la courbe de concentration intragroupe $C(P)$.

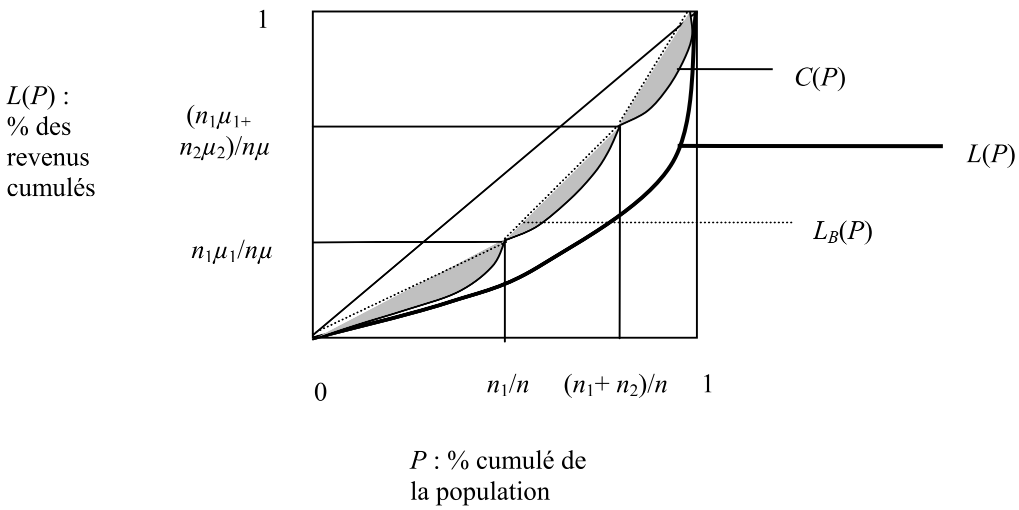
3. *Les chevauchements entre les distributions de revenu des différents groupes*. Les individus sont classés par groupe. De cette façon, le plus pauvre du groupe 2 peut se situer au-dessus d'un individu plus riche mais se situant dans le groupe 1 (où la moyenne est plus faible). Il faut donc reclasser les individus pour éviter ces chevauchements.

En définitive, on commence par calculer les inégalités intergroupes (basées sur les moyennes des groupes). Les inégalités intragroupes sont ensuite mesurées, en négligeant les effets de chevauchement. La « vraie » courbe de Lorenz est ainsi calculée. De cette manière, on obtient la procédure suivante :

$$P \to L_B(P) \to C(P) \to L(P) \,. \tag{57}$$

Illustrons ces phases successives en prenant l'exemple de trois groupes.

Figure 2 : Etapes de la décomposition



La courbe de Lorenz intergroupe $L_B(P)$ permet la mesure des inégalités entre les groupes. Cette composante se calcule de la même manière que l'indicateur de Gini global. Il faut sommer les segments entre la courbe $L_B(P)$ et la bissectrice :

$$G_B = 2\int_0^1 \left(P - L_B(P)\right) dP \,. \tag{58}$$

La deuxième étape concerne les inégalités intragroupes. Elles sont définies par l'aire entre la courbe $L_B(P)$ et la courbe $C(P)$:

$$G_W = 2\int_0^1 \left(L_B(P) - C(P)\right) dP = \sum_{h=1}^{k} a_h G_h \,. \tag{59}$$

La troisième étape permet de déduire la composante de chevauchement entre les distributions :

$$R = 2\int_0^1 \left(C(P) - L(P)\right) dP \,. \tag{60}$$

Le coefficient de Gini mesuré sur la population globale est alors donné par :

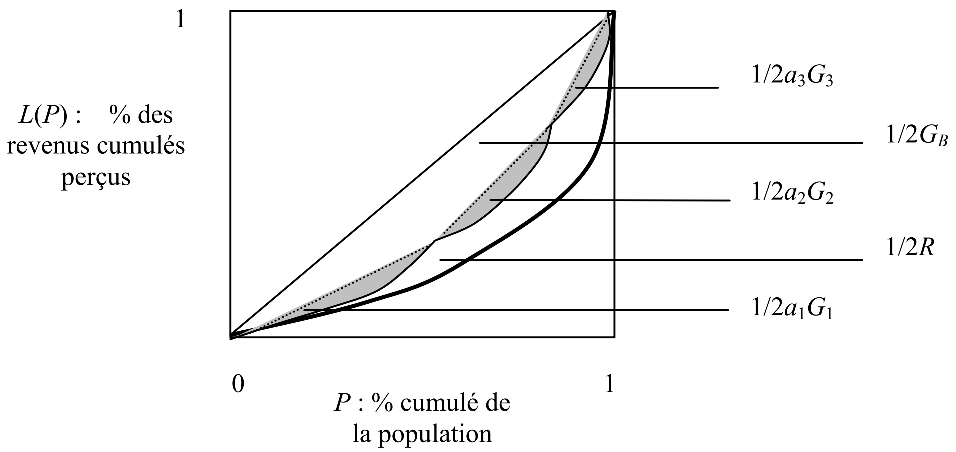
$$G = G_W + G_B + R \,. \tag{61}$$

Si la population mère est composée de trois groupes alors la décomposition s'écrit :

$$G = a_1G_1 + a_2G_2 + a_3G_3 + G_B + R \ . \tag{62}$$

La figure 3 ci-dessous permet de localiser les différents éléments de la décomposition.

Figure 3 : Localisation des aires des 3 éléments



Le terme R doit être positif. Il révèle l'intensité avec laquelle il faut permuter les individus pour atteindre la troisième étape. Il représente à la fois un terme intragroupe et un terme intergroupe : c'est un élément intergroupe à cause de son interprétation en terme de chevauchement entre les groupes et un élément intragroupe dû au chevauchement issu des inégalités intragroupes. Il est donc possible de nommer le terme R conformément à l'appellation de Mishra et Parikh (1991) : les inégalités « à travers les groupes ». Cependant, selon Nygard et Sandström (1981), la composante à travers les groupes est synthétisée par $G_B + R$. Le terme « inégalités intergroupes » est réservé à l'entropie généralisée (aux mesures additivement décomposables).

Les recherches de Lambert et Aronson défendent, au même titre que celles de Lerman et Yitzhaki (1991), la troisième composante de la décomposition du coefficient de Gini. Le terme d'interaction n'est donc pas impossible à interpréter comme le font remarquer Mookherjee et Shorrocks (1982). Lambert et Aronson dotent l'indicateur d'une information statistique essentielle et réconcilient la décomposition de l'indicateur de Gini avec les travaux de Bhattacharya et Mahalanobis (1967), puisque ces derniers montrent l'impossibilité d'unir les inégalités intragroupes avec une courbe de concentration intragroupe.

II.7. L'approche matricielle de Silber (1989) :

En poursuivant la décomposition de Pyatt (1976), Silber (1989) propose une nouvelle décomposition de l'indice de Gini. Il montre tout d'abord qu'à partir de données individuelles, l'indicateur s'exprime par :

$$G = e' G^* s , \tag{69}$$

où e' est un vecteur ligne dont les éléments sont égaux à $(1/n)$, s est un vecteur colonne dont les éléments (classés par ordre décroissant) sont les parts de revenu total détenues par chaque individu et G^* une matrice $(n \times n)$. La diagonale principale de G^* est constituée de 0. Les éléments sous la diagonale sont égaux à 1 et les éléments au-dessus de la diagonale à –1. Les individus sont répartis en sous-groupes. L'indice de Gini intergroupe est alors défini par :

$$G_B = f' G^* s , \tag{70}$$

où f' est le vecteur ligne des proportions de population globale de chaque classe de revenu et s le vecteur colonne des proportions de revenu de chaque classe. Les vecteurs f et s sont ordonnés par rapport au classement décroissant des ratios s_j/f_j (où $j = 1,\ldots, k$ est le sous-indice représentant les différents groupes). La contribution à l'inégalité totale des inégalités qui s'exercent à l'intérieur des groupes est définie par :

$$G_W = \sum_{j=1}^{k} f_j s_j G_{jj} . \tag{71}$$

L'indice de Gini G_{jj} mesuré sur les revenus du groupe j est :

$$G_B = G_{jj} = e' G^* s_{ij} , \tag{72}$$

où e' et s_{ij} sont respectivement les vecteurs de proportions des individus du groupe j (dans la population totale) et la proportion de revenu de chaque individu du groupe j. La décomposition de l'indicateur de Gini de Silber (1989) est donc :

$$G = G_W + G_B + G_O , \tag{73}$$

où G_O est le terme résiduel mesurant l'intensité de chevauchement entre les distributions. Cette intensité de chevauchement sera ensuite reprise et spécifiée par Dagum (1997a, 1997b) qui s'inspire des travaux de Gini lui-même (1916).

II.8. L'approche de C. Dagum (1997a, 1997b). Une spécification des trois composantes en cohésion avec les travaux précurseurs de Gini : la transvariation :

Une population mère P, où prévalent n unités de revenu x_i (i=1,…, n), est partitionnée en k sous-populations P_j (j=1,…, k) où P_j est de taille n_j, de fonction de répartition $F_j(x)$ et de moyenne μ_j.

On note $F(x)$ et μ, respectivement, la fonction de répartition et la moyenne mesurées sur P. Pour faire apparaître les revenus des k sous-populations, le vecteur de revenu sur P s'écrit :

$$\big((x_1,..., x_i,..., x_n) = (x_{11},..., x_{1n1}),..., (x_{j1},..., x_{jnj}),..., (x_{k1},..., x_{knk})\big). \tag{74}$$

A partir du vecteur des revenus, le coefficient de Gini mesuré sur P est donné par :

$$G = \frac{\sum_{i=1}^{n}\sum_{r=1}^{n}|x_i - x_r|}{2n^2\mu} \ . \tag{75}$$

L'indice de Gini associé à la sous-population P_j est :

$$G_{jj} = \frac{\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}|x_i - x_r|}{2n_j^2\mu_j} \ . \tag{76}$$

La différence moyenne des revenus Δ_{jh} est une généralisation de la différence moyenne de Gini (nommée aussi *GMD*). Elle représente la moyenne des différences de revenu des $n_j \times n_h$ combinaisons binaires des individus appartenant à P_j et P_h :

$$\Delta_{jh} = \frac{\sum_{i=1}^{n_j}\sum_{r=1}^{n_h}|x_{ji} - x_{hr}|}{n_j \, n_h} \ . \tag{77}$$

Elle mesure la différence de revenu espérée entre un individu tiré au hasard de la sous-population j et un individu tiré au hasard de la sous-population h.[25] Δ_{jh} permet de calculer le coefficient de Gini entre deux groupes [G_{jh}, Dagum (1987)] s'exprimant sous la forme :

$$G_{jh} = \frac{\Delta_{jh}}{\mu_j + \mu_h} \ , \ \forall\, j, h = 1,..., k \ . \tag{78}$$

[25] On retrouve un des aspects essentiels des travaux de Pyatt (1976) en rapport avec la théorie des jeux et le gain espéré que chaque individu peut gagner.

Lorsque les groupes P_j et P_h sont de même taille et distribués de manière identique, l'indice de Gini mesuré entre deux groupes est égal à l'indice de Gini intragroupe mesuré sur P_j ou P_h :

$$G_{jj} = G_{hh} = \frac{\Delta_{jj}}{2\mu_j}\ ,\ j = h = 1,....,k\ . \tag{79}$$

Remarquons aussi que les indicateurs associés à deux groupes sont symétriques :

$$G_{jh} = G_{hj} \ \ et\ \ \Delta_{jh} = \Delta_{hj}\ . \tag{80}$$

La différence moyenne de Gini Δ_{jh} mesurée entre les sous-populations P_j et P_h peut se réécrire à partir de deux concepts.

- Le premier est *la distance directionnelle brute* d_{jh}. Il s'agit d'une moyenne pondérée des différences de revenu $x_{ji}–x_{hr}$ pour chaque revenu x_{ji} d'un membre de P_j supérieur au revenu x_{hr} d'un membre de P_h, étant donné qu'en moyenne le groupe P_j est plus riche que le groupe P_h. La distance directionnelle brute se formule de la manière suivante :

$$d_{jh} = \int_0^{\infty} dF_j(x)\int_0^{x}(x-y)\,dF_h(y)\ . \tag{81}$$

- Le deuxième concept est *le moment d'ordre* 1 *de transvariation* p_{jh} entre la $j^{\text{ème}}$ et la $h^{\text{ème}}$ sous-population avec $\mu_j > \mu_h$. Il s'agit de la moyenne pondérée des différences de revenu $x_{hr}–x_{ji}$ pour chaque revenu x_{hr} d'un membre de P_h plus important que le revenu x_{ji} d'un membre de P_j. L'expression transvariation [Gini (1916), Dagum (1959, 1960, 1961)] désigne les différences de revenu qui sont de signe opposé à celui de la différence des moyennes des sous-groupes correspondants. Elle est donnée par :

$$p_{jh} = \int_0^{\infty} dF_h(x)\int_0^{x}(x-y)\,dF_j(y)\ . \tag{82}$$

Ces deux concepts sont liés par la relation suivante :

$$d_{jh} + p_{jh} = \Delta_{jh}\ . \tag{83}$$

On démontre que lorsque deux distributions ne se chevauchent pas, le moment d'ordre 1 de transvariation est nul : $p_{jh} = 0 \Rightarrow d_{jh} = \Delta_{jh}$. Aussi, lorsque la distance économique brute est égale au moment d'ordre 1 de transvariation les moyennes des deux distributions sont égales : $p_{jh} = d_{jh} = ½\ \Delta_{jh} \Rightarrow \mu_j = \mu_h$.

Dagum (1997a, 1997b) conclut d'après les résultats précédents que :

$$0 \le p_{jh} \le \frac{1}{2}\Delta_{jh} \le d_{jh} \le \Delta_{jh}\ . \tag{84}$$

La distance économique brute et le moment d'ordre 1 de transvariation permettent de définir *la richesse économique nette* entre les sous-populations P_j et P_h : $d_{jh} - p_{jh}$ $\forall \mu_j > \mu_h$. Etant donné que $d_{jh} - p_{jh}$ admet 0 comme origine et Δ_{jh} comme maximum, on peut normaliser la richesse économique nette, qui devient ainsi la richesse économique relative ou encore distance économique directionnelle.[26]

- *La richesse économique relative* D_{jh} (ou distance économique directionnelle) entre les sous-populations P_j et P_h est le ratio entre la richesse économique nette et son maximum Δ_{jh} :

$$D_{jh} = \frac{d_{jh} - p_{jh}}{\Delta_{jh}} = \frac{d_{jh} - p_{jh}}{d_{jh} + p_{jh}} \ . \tag{85}$$

D_{jh} est inclus dans l'intervalle fermé [0,1]. C'est un nombre sans dimension car d_{jh}, p_{jh} et Δ_{jh} ont la dimension du revenu. La richesse économique relative sépare les inégalités intergroupes en deux composantes :

- la contribution nette des inégalités de revenu entre les sous-populations P_j et P_h, obtenue par le produit $G_{jh} \times D_{jh}$;

- la contribution des intensités de transvariations entre P_j et P_h, obtenue par $G_{jh} \times (1 - D_{jh})$. L'addition de ces deux produits mesure de manière brute les inégalités de revenu entre deux sous-populations P_j et P_h.

La décomposition du coefficient de Gini, proposée par Dagum (1997a, 1997b) est séparable en trois éléments. Ce sont des contributions distinctes de l'inégalité globale mesurée sur la population mère P. Dagum (1997a) montre que l'indice de Gini total calculé sur P s'écrit :

$$G = \frac{\sum_{j=1}^{k} \sum_{h=1}^{k} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} \left| x_{ji} - x_{hr} \right|}{2n^2 \mu} \ . \tag{86}$$

Cette expression fait apparaître les différences de revenu intragroupes et intergroupes. Les caractéristiques des sous-groupes sont très importantes. Elles sont décisives quant à l'évaluation de la contribution de chaque groupe à l'inégalité totale.

[26] La littérature sur les distances entre les distributions s'est développée notamment avec les travaux de Dagum (1980), Shorrocks (1982), Ebert (1984), Chakravarty et Dutta (1987), Yitzhaki (1994), et Deutsch-Silber (1997). Il est aussi possible de s'intéresser à la critique de Shorrocks (1982) concernant la mesure de Dagum (1980) et à la réponse de Dagum (1997a, p. 519). Shorrocks (1982) montre que la mesure de Dagum ne reflète pas une distance euclidienne. Or, Dagum (1997a) fait remarquer à Shorrocks qu'il s'agit d'une distance économique directionnelle conformément à la définition de 1980, conduisant Shorrocks à une erreur d'interprétation fondamentale [voir aussi la critique de Yitzhaki (1994) sur la distance D_1 de Dagum].

Ces spécificités sont notamment le pourcentage d'individus appartenant au groupe P_j (p_j) et le pourcentage de revenus de P_j lié au revenu global de la population P (s_j) :

$$p_j = \frac{n_j}{n}, \; s_j = \frac{n_j \, \mu_j}{n \, \mu} \; . \tag{87}$$

D'après ces pondérations, l'indice de Gini calculé sur P devient :

$$G = p' \, \Phi \, s \, . \tag{88}$$

Dans cette équation, Φ est une matrice symétrique ($k{\times}k$) où les éléments sont les indices de Gini entre l'ensemble des paires de sous-populations P_j et P_h (G_{jh}). Les éléments de Φ situés sur la diagonale principale sont les indices de Gini associés aux groupes P_j (G_{jj}, $\forall j = 1,\ldots, k$). Les vecteurs p et s sont les proportions de population et de revenu des k sous-populations (dont les éléments sont respectivement les p_j et s_j). En développant l'équation, on met en évidence, les termes de la matrice Φ :

$$G = \sum_{j=1}^{k} G_{jj} \, p_j \, s_j + \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(p_j \, s_h + p_h \, s_j \right) . \tag{89}$$

Deux sommes apparaissent. La première indique la contribution des inégalités à l'intérieur des k sous-groupes. La deuxième composante, la double somme, est la contribution des inégalités entre les C_k^2 paires de sous-populations. Elle représente les inégalités intergroupes brutes :

$$G_{gb} = G_{nb} + G_t = \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(p_j \, s_h + p_h \, s_j \right) . \tag{90}$$

Comparées aux inégalités moyennes entre les groupes, les inégalités intergroupes brutes offrent davantage d'informations puisqu'elles mettent en évidence les différences de revenu entre chaque paire de groupes. Ensuite, en multipliant la contribution des inégalités intergroupes par D_{jh} puis par $1–D_{jh}$ (soit au total par 1), nous pouvons distinguer les inégalités intergroupes nettes [G_{nb}] des inégalités intergroupes de transvariation [G_t] :

$$G = \sum_{j=1}^{k} G_{jj} \, p_j \, s_j + \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \, D_{jh} \left(p_j \, s_h + p_h \, s_j \right) + \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(1 - D_{jh} \right) \left(p_j \, s_h + p_h \, s_j \right) . \tag{91}$$

La contribution des inégalités de revenu inhérentes à l'intensité de la transvariation,

$$G_t = \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(1 - D_{jh} \right) \left(p_j \, s_h + p_h \, s_j \right) , \tag{92}$$

mesure le poids des inégalités intergroupes issues du chevauchement entre les distributions. Il s'agit d'inégalités particulières. Le chevauchement signifie que certains individus de la distribution la plus pauvre possèdent des revenus supérieurs aux personnes de la distribution la plus riche. L'intensité de la transvariation permet de capter les inégalités générées par les hauts revenus des sous-populations les plus pauvres. A contrario, les disparités provenant des revenus élevés des sous-populations riches sont données par la contribution nette des inégalités intergroupes à l'inégalité totale :

$$G_{nb} = \sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\, D_{jh}\left(p_j\, s_h + p_h\, s_j\right). \tag{93}$$

Il s'agit des inégalités entre les k sous-populations dont les revenus sont issus de la partie de non-chevauchement entre les distributions. Il s'agit aussi des inégalités moyennes entre les groupes puisque : $\mu_j = \mu_h$ ($\forall j,\ h = 1,\ldots, k$) $\Rightarrow G_{nb} = 0$. Le dernier élément de la décomposition est la contribution des inégalités intragroupes à l'inégalité totale :

$$G_w = \sum_{j=1}^{k} G_{jj}\, p_j\, s_j\ . \tag{94}$$

Il s'agit d'une simple moyenne pondérée des indicateurs de Gini associés aux groupes P_j.

Les trois contributions de l'inégalité globale permettent d'introduire l'équation fondamentale de la décomposition de l'indicateur de Gini[27],

$$\boxed{G = G_w + G_{nb} + G_t\ ,} \tag{95}$$

où :

- G_w est la contribution des inégalités à l'intérieur des sous-populations ;
- G_{nb} est la contribution nette des inégalités entre les sous-populations ;
- et G_t l'inégalité inhérente à l'intensité de transvariation entre les sous-populations.

Les inégalités sont donc partagées en contributions intragroupes et intergroupes. La partie intergroupe est conditionnée par le comportement des distributions. Pour des distributions de même moyenne, la composante intergroupe est uniquement constituée des transvariations intergroupes :

$$G_{nb} = 0,\ G_w = \sum_{j=1}^{k} G_{jj}\, p_j^{\,2},\ G_t = 2\sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\, p_j\, p_h \ \ ssi\ \mu_1 = \ldots = \mu_k = \mu\,. \tag{96}$$

Si les distributions sont indépendantes et identiquement distribuées (*i.i.d.*), on a :

$$D_{jh} = 0,\ G_{nb} = 0,\ G_{jh} = G_{jj} = G = G_w = G_{nb} = G_t = 0\ . \tag{97}$$

[27] Pour toutes programmations informatiques, confer Dagum, Mussard, Seyte et Terraza (2002).

En revanche, si les k distributions ne se chevauchent pas alors les distances économiques entre chaque groupe sont égales à l'unité :

$$D_{jh} = 1 \;\Leftrightarrow\; G_t = 0 \;\; (\forall j, h = 1, \ldots, k). \tag{98}$$

En définitive, la décomposition du coefficient de Gini permet de prendre connaissance des groupes générateurs d'inégalité et de savoir dans quelles mesures ils s'éloignent des autres sous-populations dans le souci d'atteindre une répartition plus égalitaire des revenus, préalable nécessaire à la redistribution entre les groupes.

III. Conclusion :

On a longtemps pensé que l'indice de Gini n'était pas une mesure décomposable en trois éléments. En effet, les indicateurs étaient jusque dans les années 80 tous décomposables en deux termes. Il était difficile, en outre, de comprendre le rôle joué par le troisième élément (G_t). Celui-ci a été comparé à un résidu ou un terme d'interaction [voir Mookherjee et Shorrocks (1982)], jusqu'aux travaux novateurs de Silber (1989), Deutsch et Silber (1997) et Dagum (1997a, 1997b). En particulier, en 1997, ils ont été les premiers (dans la même revue) à démontrer, de manière différente, que le troisième élément est issu d'un concept introduit par Gini (1916) et repris par Dagum (1959, 1960, 1961) : la transvariation. Il s'agit d'un changement dans le domaine de la décomposition de l'indicateur de Gini. En effet, ces travaux ont permis de réconcilier l'indice de Gini avec une décomposabilité (en deux ou trois éléments) dont chaque contribution est un élément clair et défini de l'inégalité totale, et spécifié en cohésion avec les travaux précurseurs de Gini.

Les premières limites ont été soulevées par Shorrocks (1980, 1984) avec la notion de cohérence en sous-groupes [voir aussi Shorrocks (1988) et Cowell (1988)]. Les critiques de Shorrocks et Cowell s'appuient sur le fait que leur mesure (la classe de l'entropie généralisée) satisfait la décomposabilité additive et la cohérence en sous-groupes. Adelman et Levy (1984, 1986) tentent de démontrer à Cowell (1985) que la mesure de Theil ne satisfait pas une décomposabilité non ambiguë. Même si d'un point de vue mathématique celle-ci est correcte, la mesure de Theil peut conduire à des mauvaises estimations et donc à des conclusions erronées lorsqu'on s'intéresse aux facteurs ou aux sources des revenus individuels. En effet, les composantes intragroupes et intergroupes n'intègrent pas la fonction de revenu, ce qui permet seulement de prendre en compte des effets partiels sur la mesure globale de l'inégalité. La remarque d'Adelman et Levy est intéressante car elle met en exergue les difficultés auxquelles le chercheur fait face lorsque ce dernier tente d'analyser les sources de revenu à l'aide de l'entropie, ce qui n'est pas le cas en utilisant l'indice de Gini.

La décomposition de l'indice de Gini de 1997 (C.Dagum) permet d'identifier :

- les inégalités entre les groupes ;
- et les inégalités à l'intérieur des groupes qui composent la population.

Les inégalités à l'intérieur des groupes sont représentées par G_w. et les inégalités existantes entre les groupes sont symbolisées par G_{gb}. L'indice de Gini se décompose ainsi :

$$G = G_{gb} + G_w.$$

Contrairement aux autres indices, comme ceux issus de l'entropie, la composante intergroupe brute G_{gb} n'est pas assimilable à une inégalité moyenne entre les groupes. Il s'agit d'un indice qui mesure les différences entre toutes les paires de revenu provenant d'un groupe différent. On comprend alors que la propriété de comparaisons interpersonnelles [déjà présente dans Pyatt (1976)] est une propriété fondamentale de décomposition. En effet, en regroupant les comparaisons issues de chaque groupe et celles émanant de groupes différents, on met en évidence les inégalités intragroupes et les inégalités (brutes) intergroupes. Par conséquent, contrairement aux indices issus de la seconde loi de l'entropie, la décomposition de l'indice de Gini repose sur des fondements normatifs plus complets (les comparaisons interpersonnelles de revenu ou d'utilité sont des propriétés économiques incontournables).

La mesure brute intergroupe offre par ailleurs une vision duale des disparités entre deux groupes, permettant une décomposition en trois termes. La dualité de l'élément intergroupe s'exprime d'une part, par les différences de revenu moyennes entre les groupes (G_{nb}) et d'autre part, par les inégalités issues du chevauchement entre les distributions (G_t). La décomposition de l'indice de Gini permet donc de mesurer :

- la contribution des inégalités intragroupes (G_w) à l'inégalité totale ;
- la contribution des inégalités intergroupes nettes (les inégalités moyennes intergroupes, G_{nb}) ;
- et la contribution de l'intensité de transvariation (G_t) représentant les inégalités intergroupes générées par les revenus élevés des groupes relativement plus pauvres (par rapport à la moyenne des revenus) provenant du chevauchement entre les distributions.

Cette particularité intergroupe concède à l'indicateur de Gini une structure singulière qui peut inciter les chercheurs à recourir à son utilisation.

BIBLIOGRAPHIE

Adelman I. et Levy A. (1984), « Decomposing Theil's Income of Income Inequality into Between and Within Components : A Note », *Review of Income and Wealth*, Vol. 30 : p. 119-121.

Adelman I. et Levy A. (1986), « Decomposing Theil's Index of Income Inequality : A Reply », *Review of Income and Wealth*, Vol. 32 : p. 107-108.

Atkinson A. (1970), « On the Measurement of Inequality », *Journal of Economic Theory*, Vol. 55 : p. 244-263.

Bhattacharya N. et Mahalanobis B. (1967), « Regional Disparities in Household Consumption in India », *Journal of the American Statistical Association*, Vol. 62 : p. 143-161.

Blackorby C., Bossert W. et Donaldson D. (1999), « Income Inequality Measurement : The Normative Approach », dans Silber J. (ed.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, p.133-161.

Bossert W. et Pfingsten A. (1990), « Intermediate Inequality : Concepts, Indices and Welfare Implication. », *Mathematical Social Science*, Vol. 19 : p. 117-134.

Bourguignon F. (1979), « Decomposable Inequality Measures », *Econometrica*, Vol. 47 : p. 901-920.

Chakravarty S.R. (1999), « Measuring Inequality : The Axiomatic Approach », dans Silber J. (ed.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, p.163-186.

Chakravarty S.R. et Dutta B. (1987), « A Note on Measures of Distance Between Income Distributions », *Journal of Economic Theory*, Vol. 41 : p. 185-188.

Champernowne G. (1952), « The Graduation of Income Distribution », *Econometrica*, Vol. 20(4) : p. 591-615.

Cowell F. A. (1980a), « Generalized entropy and the Measurement of Distributional Change », *European Economics Review*, Vol. 13 : p. 147-159.

Cowell F. A. (1980b), « On the Structure of Additive Inequality Measures », *Review of Economics Studies*, Vol. 47 : p. 521-531.

Cowell F. A. (1988), « Inequality Decomposition : Three Bad Measures », *Bulletin of Economic Research*, Vol. 40(4) : p. 309-312.

Cowell F. A. (2000), « Measurement of Inequality », dans Atkinson A. et Bourguignon F. (ed.), *Handbook of Income Distribution*.

Cowell F. A. et Kuga K. (1981), « Additivity and the Entropy Concept : An Axiomatic Approach to Inequality Measurement », *Journal of Economic Theory*, Vol. 25 : p. 131-143.

Dagum C. (1959), *Transvariazione fra più di due distribuzioni*, *In : Gini, C.(ed.) Memorie di metodologia statistica*, Vol II, Libreria Goliardica, Roma.

Dagum C. (1960), « Teoria de la transvariacion, sus aplicaciones a la economia », *Metron*, Vol. XX : p. 1-206.

Dagum C. (1961), « Transvariacion en la hipotesis de variables aleatorias normales multidimensionales », *Proceedings of the International Statistical Institute*, Vol. 38, Book 4, Tokyo, p. 473-486.

Dagum C. (1977), « A New Model of Personal Income Distribution : Specification and Estimation », *Economie Appliquée*, Vol XXX (3) : p. 413-436.

Dagum C. (1980), « Inequality Measures Between Income Distributions with Applications », *Econometrica*, Vol. 48 (7) : p. 1791-1803.

Dagum C. (1987), « Measuring the Economic Affluence Between Populations of Income Receivers », *Journal of Business and Economic Statistics*, Vol. 5(1) : p. 5-12.

Dagum C. (1997a), « A New Approach to the Decomposition of the Gini Income Inequality Ratio », *Empirical Economics*, Vol. 22(4) : p.515-531.

Dagum C. (1997b), « Decomposition and Interpretation of Gini and the Generalized Entropy Inequality Measures », *Proceedings of the American Statistical Association, Business and Economic Statistics Section, 157th Meeting,* p. 200-205.

Dagum C. (1998), « Fondements de bien-être social et décomposition des mesures d'inégalité dans la répartition du revenu », *Economie Appliquée*, Tome L1 n°4 : p. 151-202.

Dagum C. (1999), « Linking the Functional and Personal Distributions of Income », dans Silber J. (ed.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, p.101-128.

Dagum C., Mussard S., Seyte F. et Terraza M. (2002), *Programme pour la décomposition de l'indice de Gini de C. Dagum*, disponible sur <http://www.lameta.univ-montp1.fr/online/gini.html>.

Dalton H. (1920), « The Measurement of Inequality of Incomes », *Economic Journal*, Vol. 30 : p. 348-361.

Das T. et Parikh H. (1982), « Decomposition of Inequality Measures and a Comparative Analysis », *Empirical Economics*, Vol. 7 : p. 23-48.

Deutsch J. et Silber J. (1997), « Gini's 'Transvariazione' and the Measurement of Distance Between Distributions », *Empirical Economics*, Vol. 22 : p. 547-554.

Deutsch J. et Silber J. (1999), « Inequality Decomposition by Population Subgroups and the Analysis of Interdistributional Inequality », dans Silber J. (ed.), *Handbook of Income Inequality Measurement*, Kluwer Academic Publishers, p.163-186.

Ebert U., (1984), « Measures of Distance Between Income Distributions », *Journal of Economic Theory*, Vol. 32 : p. 266-274.

Ebert U., (1988), « On the Decomposition of Inequality : Partitions into Non-overlapping Subgroups », dans Eichorn W. (ed.), *Measurement In Economics*. New-York : Physica Verlag, p. 399-412.

Foster J. E. (1983), « An Axiomatic Characterization of the Theil Measure of Income Inequality », *Journal of Economic Theory*, Vol. 31 : p. 105-121.

Foster J. E., Greer J. et Thorbecke E. (1984), « Notes and Comments. A Class of Decomposable Poverty Measures », *Econometrica*, Vol. 52 : p. 761-766.

Foster J. E. et Schneyerov A. A. (1997), « Path Independent Inequality Measures », Working Paper 96-W04 (révisé), Département d'Economie, Vanderbilt University.

Gini C. (1916), « Il concetto di transvariazione e le sue prime applicazioni », *Giornale degli Economisti e Rivista di Statistica*, dans Gini C. (ed.) (1959) : p.21-44.

Gini C. (1921), « Measurement of Inequality of Incomes », *The Economic Journal*, March : p. 124-126.

Hart P.E. (1970), « Entropy and Others Measures of Concentration », *Journal of the Royal Statistical Society*, Séries A 134 : p. 73-85.

Kolm S-C. (1966), « The Optimal Production of Social Justice », *Colloques Internationaux du CNRS*, Biarritz, 2-9 septembre 1966.

Kolm S-C. (1976a), « Unequal Inequalities I », *Journal of Economic Theory*, Vol. 12 : p. 416-442.

Kolm S-C. (1976b), « Unequal Inequalities II », *Journal of Economic Theory*, Vol. 13 : p. 82-111.

Kolm S-C. (1977), « Multidimensional Egalitarianisms », *Quarterly Journal of Economics*, Vol XCI n°1 : p. 1-13.

Lambert P. J. et Aronson R. J. (1993), « Inequality Decomposition Analysis and the Gini Coefficient Revisited », *The Economic Journal*, Vol. 103 : p. 1221-1227.

Lerman R. et Yitzhaki S. (1984), « A Note on the Calculation and Interpretation of the Gini Index », *Economics Letters*, Vol. 15 : p.363-368.

Lerman R. et S. Yitzhaki (1991), « Income Stratification and Income Inequality », *Review of Income and Wealth*, Vol. 37(3) : p. 313-329.

Lorenz M. O. (1905), « Methods for Measuring Concentration of Wealth », *Journal of the American Statistical Association*, Vol. 70 : p. 209-219.

Mangahas M. (1975), « Income Inequality in the Philippines : a Decomposition Analysis », World Employement Programme Working Paper N°12, Genève, Office international du travail.

Mesnard (de) L. (1997), « A propos des problèmes causés par les indices de mesure d'inégalité de Gini et de Kakwani », *Document de travail n° 9713*, Université de Bourgogne.

Milanovic B. et Yitzhaki S. (2002), « Decomposing World Income Distribution: Does the World Have a Middle Class? », *Review of Income and Wealth*, Vol. 48(2) : p. 155-78.

Mishra V. et Parikh A. (1991), « Houshold Consummer Expenditure Inequalites in India : a Decomposition » , Economics Discussion Paper N° 9111, University of East Anglia.

Mookherjee D. et A. Shorrocks (1982), "A Decomposition Analysis of the Trend in UK Income Inequality", *The Economic Journal*, Vol. 92 : p. 886-902.

Mussard S., Seyte F. et Terraza M. (2003a), « Decomposition of Gini and the Generalized Entropy Inequality Measures », *Economics Bulletin*, Vol.4 n°7 : p. 1-6.

Nygard et Sandström (1981), *Measuring Income Inequality*, Stockholm : Almqvist et Wicksell.

Pareto V. (1896), « *Ecrits sur la courbe de la répartition de la richesse*. Œuvres complètes de Vilfredo Pareto publiées sous la direction de Giovanni Busino. Genève : Librairie Droz, 1965.

Pyatt G. (1976), « On the Interpretation and Disaggregation of Gini Coefficients », *Economic Journal*, Vol. 86 : p. 243-25.

Pyatt G., Chen C-N, Fei J. (1980), « The Distribution of Income by Factor Component », *Quarterly Journal of Economics*, Vol. 95(3) : p. 451-473.

Rao V.M. (1969), « Two Decompositions of Concentration Ratio », *Journal of the Royal Statistical Society*, Séries A 132 : p. 418-425.

Russel R. R. (1985), « A Note on Decomposable Inequality Measures », *Review of Economic Studies*, Vol. LII : p. 347-352.

Salas R. (2002), « Multilevel Interterritorial Convergence and Additive Multidimensional Inequality Decomposition », *Social Choice and Welfare*, Vol. 19 : p. 207-218.

Sen A. K. (1973), *On Economic Inequality*, Clarendon Press, Oxford.

Shorrocks A. F. (1980), « The Class of Additively Decomposable Inequality Measures », *Econometrica*, Vol. 48 : p. 613-625.

Shorrocks A. F. (1982), « On the Distance Between Income Distributions », *Econometrica*, Vol. 50 : p. 1337-1339.

Shorrocks A. F. (1984), « Inequality Decomposition by Population Subgroups », *Econometrica*, Vol. 52 : p. 1369-1386.

Shorrocks A. F. (1988), « Aggregation Issues in Inequality Measurement », In W. Heichhorn, ed., *Measurement in Economics*, New York, Physica-Verlag, p. 429-452.

Silber J. (1989), « Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality », *Review of Economics and Statistics*, Vol. 71 : p. 107-115.

Singh S. K. et Maddala G. S. (1976), « A Function for Size Distribution of Income », *Econometrica*, Vol. 25(4) : p. 568-590.

Soltow L. (1960), « The Distribution of Income Related to Changes in the Distribution of Education, Age and Occupation », *Review of Economics and Statistics*, Vol. 42 : p. 450-453.

Theil H. (1967), *Economics and Information Theory*, North-Holland Publishing Company, Amsterdam.

Wagstaff A., (2003), "Inequality Aversion, Health Inequalities and Health Achievement", *Journal of Health Economics*, Vol. 19 : p. 627-641.

Yitzhaki S. (1994), « Economic Distance and Overlapping of Distributions », *Journal of Econometrics*, Vol. 61 : p. 147-159.