

UNIVERSITÉ PARIS DAUPHINE



N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

Pour l'obtention du titre de
DOCTEUR EN INFORMATIQUE
(Arrêté du 7 août 2006)

Spécialité : Aide à la Décision

<p>CONSTRUCTION D'INDICES : MESURAGE MULTIDIMENSIONNEL DE LA PAUVRETÉ</p>
--

Candidat : Vivien KANA ZEUMO

JURY

- Directeurs de thèse : Alexis TSOUKIÀS
Directeur de Recherche CNRS – Université Paris Dauphine
Blaise SOMÉ
Professeur Titulaire à l'Université de Ouagadougou
- Rapporteurs : Marc PIRLOT
Professeur à la Faculté Polytechnique de Mons
Ivan BLECIC
Professeur à l'Université de Sassari
- Examineurs : Daniella GRIGORI
Professeur à Université Paris Dauphine
Alberto COLORNI
Professeur à Politecnico di Milano

Présentée et soutenue publiquement le 18/07/2012

L'université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses : ces opinions doivent être considérées comme propres à leurs auteurs.

“Sous un bon gouvernement, la pauvreté est une honte ; sous un mauvais gouvernement, la richesse est aussi une honte.”

Confucius -551–479 avant Jésus-Christ : Livre des sentences

“Under a good government, poverty is a shame ; under a bad government, the richness is also a shame.”

Confucius -551–479 before Jesus-Christ : Book of sentences

Remerciements

Ce travail a été effectué au sein du laboratoire LAMSADE de l'Université Paris Dauphine et de l'Unité de Formation et de Recherche en Sciences exactes et appliquées (UFR–SEA) de l'Université de Ouagadougou. Quelques lignes ne pourront jamais exprimer la reconnaissance que j'éprouve envers tous ceux qui, de près ou de loin, ont contribué par leurs conseils, leurs encouragements et leur amitié à l'aboutissement de cette thèse.

Je remercie infiniment mon directeur de thèse, M. Alexis TSOUKIÀS (Directeur de Recherche CNRS), qui a accepté d'encadrer cette thèse et qui m'a témoigné son soutien et sa confiance. Il a su me faire bénéficier de son expérience et de ses compétences. Je le remercie pour le soutien, la confiance, le suivi et la qualité de son encadrement. Qu'il trouve ici l'expression de ma plus haute considération et de mes sincères gratitude.

Je remercie mon directeur de thèse, le Prof. Blaise SOMÉ (Professeur titulaire), pour son soutien et ses conseils judicieux tout au long de cette thèse. Il a su me faire bénéficier de sa confiance, de son expérience et de ses compétences. Qu'il trouve ici l'expression de ma plus haute considération et de mes sincères gratitude.

Je tiens également à remercier, le Prof. Vangelis Th. PASCHOS pour m'avoir accueilli au sein du laboratoire LAMSADE et avoir permis que ce travail se déroule dans les meilleurs conditions.

J'exprime mes sincères gratitude aux membres du jury, Prof. Ivan BLECIC (Università di Sassari), Prof. Alberto COLORNI (Politecnico di Milano), Prof. Daniella GRIGORI (Université Paris Dauphine), Prof. Marc PIRLOT (Faculté Polytechnique de Mons) pour l'honneur qu'ils m'ont fait en acceptant de juger cette thèse ou d'en être les rapporteurs.

Je remercie les professeurs Marc PIRLOT et Denis BOUYSSOU pour l'honneur qu'ils m'ont fait en acceptant d'être rapporteurs ce travail de recherche lors de ma pré-soutenance. J'exprime ma reconnaissance aux membres du jury de ma pré-soutenance, Dr. Hadiza MOUSSA-SALEY et Dr. Longin SOMÉ. Leurs conseils, remarques et critiques m'ont permis d'améliorer considérablement ce document.

Je tiens à remercier Stefano MORETTI pour tous ses conseils et remarques.

Je ne saurais oublier mes collègues de laboratoire pour leur cordialité grâce à laquelle cette thèse s'est déroulée dans un climat stimulant et sympathique. Je tiens

à remercier : Renaud, Olivier, Katerina, Mireille, Amidou, Guilia, Sonia, Élisée, Wassila, Renaud, Valerie.

Je tiens à remercier tout particulièrement les professeurs Mel Janowitz (Rutgers University), David Ríos INSUA (Universidad Rey Juan Carlos), Denis BOUYS-SOU, Edwin DIDAY, Bernard ROY pour leurs remarques, leurs conseils et encouragements durant la préparation de cette thèse.

J'aimerais remercier tout autant mes ami(e)s qui ont contribué à leur manière à l'aboutissement de cette thèse.

J'ai passé plus de six mois au LAMSADE dans le bureau de Meltem OZTURK et Stefano MORETTI. Je les en remercie tout deux.

Je n'oublierai jamais le rôle de ma famille (mon père, ma mère, mes frères et ma sœur) qui a été capital tout au long de ces années.

L'aide financière de mon grand-frère Yves N. KANA m'a été très utile pour mener à bien ce travail de recherche. Qu'il trouve ici mes sincères remerciements.

Je remercie Adèle ZONGO pour son soutien et ses encouragements durant la préparation de cette thèse. Qu'elle trouve ici mes sincères gratitude.

Merci enfin au DIEU tout Puissant refuge, espoir, richesse et puissance des pauvres ; au DIEU Amour et seul feu capable de consumer l'égoïsme, racine de la pauvreté.

Il se peut que j'ai oublié quelques personnes qui m'ont aidé toute au long de ce travail de recherche. Qu'ils retrouvent ici mes meilleures reconnaissances et mes sincères et profondes excuses.

Dédicaces

Je dédie ce travail à :

À mon père, Étienne KANA, décédé pendant la préparation de cette thèse, qui aurait été heureux de voir que son fils est allé au bout de son projet.

Il a été le premier à m'encourager à aller si loin dans les études. Il m'a inculqué le goût du travail, de la rigueur et de l'honnêteté. Son soutien indéfectible tout au long de mon cycle scolaire et universitaire restera à jamais gravé dans ma mémoire. J'ai voulu le mener cette thèse à terme pour honorer sa mémoire.

À ma mère Christine ZEUGUEU.

Que ce travail puisse t'apporter le bonheur souhaité. Trouvez ici l'expression de mon indéfectible amour filial.

À Alexis TSOUKIÀS.

Tu m'as toujours soutenu, encouragé et conseillé durant toutes ces années. Trouve ici ma profonde gratitude. Merci...

À mon frère aîné Yves KANA NGAFACK.

Tu as été et reste pour moi un père car tu as mis à ma disposition tous les moyens nécessaires pour assurer le succès de ma formation. Ce travail est le fruit de tes sacrifices. Merci...

À toutes les victimes de la pauvreté.

Soyez rassurés que notre attachement à mettre sur pied un modèle fiable de lutte contre la pauvreté sera toujours déterminant.

À toute ma grande famille, à tous ceux et à toutes celles dont le nom est inscrit sur la page secrète de mon cœur.

Table des matières

Introduction générale	1
1 Préliminaire et problématique générale	5
1.1 Contexte général	5
1.1.1 La place de la pauvreté dans le monde	5
1.1.2 La complexité du problème	8
1.2 La pauvreté et son mesurage	11
1.2.1 Qu'est ce que la pauvreté ?	11
1.2.2 Pourquoi mesurer la pauvreté ?	14
1.2.3 Vers une nouvelle méthodologie de mesurage	15
1.3 Nos motivations	18
1.3.1 Les besoins d'opérationnalisation des capacités	18
1.3.2 Les besoins méthodologiques	19
1.4 Les mesures significantes	22
1.4.1 Qu'est-ce qu'une mesure significative ?	23
1.4.2 Les enjeux des mesures significantes	24
1.5 Objectif générale de la thèse	26
I État de l'art	29
2 Analyse et mesure de la pauvreté	31
2.1 Introduction	31
2.2 Un peu d'histoire	32
2.2.1 La pauvreté : origines et fondements conceptuels	32
2.2.2 Les fondements du mesurage de la pauvreté	36
2.3 Mesurage multidimensionnel de la pauvreté	40
2.3.1 Notations and définitions	40
2.3.2 Les approches traditionnelles	42
2.3.3 Les approches multidimensionnelles traditionnelles	45
2.3.4 Caractérisation axiomatique	64
2.3.5 L'approche par les capacités	70
2.3.6 Les nouveaux indices multidimensionnels de pauvreté	72

II	Construction d'indices : méthodologie générale	75
3	Méthodologie pour l'analyse et le mesurage de la pauvreté	77
3.1	Concepts de base	77
3.1.1	Commodités, fonctionnements et capacités	77
3.1.2	Notations et terminologies	80
3.2	Transformation de la table des réponses en matrice de proximité	81
3.2.1	Standardisation des variables	81
3.2.2	Détermination de la matrice de proximité	82
3.3	La classification	83
3.3.1	Sélection d'algorithmes de classification	83
3.3.2	Validation des clusters	84
3.3.3	Visualisation des clusters	85
3.3.4	Matrice modal-valuée et critère modal-valué	86
3.3.5	Exemple illustratif	87
3.4	Construction des mesures significantes	90
3.4.1	Caractérisation des modalités	90
3.4.2	Spécification des informations préférentielles	92
3.4.3	Construction des capacités	93
3.4.4	Exemple illustratif	97
3.5	Implémentation des actions concrètes	103
3.6	Discussion	104
3.7	Quelques propriétés et axiomes	106
4	Méthode de filtrage multi-attributs	111
4.1	Introduction	111
4.2	Construction des clusters diagnostics	112
4.2.1	Position du problème	112
4.2.2	Les clusters diagnostics	112
4.3	La méthode d'affectation floue	114
4.3.1	Principe concordance	114
4.3.2	Principe de discordance	117
4.3.3	Règle d'affectation	118
4.4	Exemple illustratif	119
4.4.1	Construction des clusters diagnostics D_h	119
4.4.2	Calcul des indices $A(x_i, D_h)$, $E(x_i, D_h)$ et $R(x_i, D_h)$	119
4.4.3	Calcul de indices de concordance $H(x_i, D_h)$	120
4.4.4	Calcul des indices de discordance $D(x_i, D_h)$	121
4.4.5	Calcul des indices d'appartenance $B(x_i, D_h)$	121
	Conclusion générale	123

A	Application sur la base de données ASSL 2007 du Burkina Faso	125
A.1	Présentation de la base des données	125
A.2	Application	127
A.2.1	Construction des clusters diagnostics D_h	130
A.2.2	Résultats et discussion	132
B	Les techniques de clustering	135
B.1	Introduction	135
B.2	Choix d'une mesure de ressemblance	136
B.2.1	Généralités	136
B.2.2	Mesures de ressemblance entre individus	138
B.2.3	Mesures de ressemblance entre variables	143
B.3	Les approches de clustering	147
B.3.1	Les approches hiérarchiques	148
B.3.2	Les approches par partitionnement	153
B.3.3	Les approches fondées sur la densité	154
B.3.4	Les approches fondées sur un modèle	155
B.3.5	Les approches fondées sur le quadrillage de l'espace	156
B.4	Évaluation d'une classification	156
B.4.1	Détermination du meilleur nombre de clusters	156
B.4.2	Validation des partitions	163
B.4.3	Comparaison des partitions	165
B.4.4	Caractérisation des clusters	168
	Bibliographie	173

Introduction générale

La question de lutte contre la pauvreté est la préoccupation de chaque gouvernement qui, à son niveau, doit faire des choix de politiques publiques appropriées, supposés cohérents avec les conditions de vie de ses populations. Plus précisément, se pose le problème de l'homme d'étude qui devra proposer au choix du décideur des politiques de décision et stratégies de luttes cohérentes alternatives. Il va de soit que si l'opinion publique attend du décideur qu'il émette un "*discours favorable*" par rapport à l'amélioration de leur condition de vie, et surtout, qu'il prenne des mesures nécessaires et rapides afin que leur 'souffrance' soit apaisé, il n'en reste pas moins vrai que la maîtrise des agents vecteurs de la pauvreté et des systèmes en cause lui échappe en grande partie. La pauvreté est *objectivement* insoluble, pourtant, elle progresse en se métamorphosant et gêne tout le monde. Les populations des pays sous-développés, des pays en développement et des pays développés sont de plus en plus exigeants vis-à-vis de leurs états et de la communauté internationale pour trouver une solution à ce fléau. Il faut développer des stratégies adéquates opérationnelles et légitimes visant l'amélioration de la situation des populations.

Au niveau national et international, la modélisation des stratégies de lutte contre la pauvreté, l'élaboration et l'implémentation des politiques appropriées d'amélioration des conditions de vie des populations basées sur une approche multidimensionnelle représente des enjeux majeurs pour la promotion la croissance, la lutte contre la pauvreté et l'exclusion sociale, la promotion de la redistribution efficace de la richesse, la promotion de la justice sociale et la défense des droits des plus démunis par une gestion efficace de la richesse et des ressources. C'est pourquoi deux des quatre critères d'éligibilité à l'Initiative en faveur des Pays Pauvres Très Endettés (PPTE) sont : (i) avoir parfaitement mis en œuvre des réformes et de saines politiques économiques dans le cadre de programmes soutenus par le FMI et la Banque mondiale ; (ii) avoir formulé un Document de Stratégie pour la Réduction de la Pauvreté (DSRP). Le DSRP décrit les politiques et les programmes macroéconomiques, structurels et sociaux qu'un pays mettra en œuvre pendant plusieurs années pour promouvoir la croissance et réduire la pauvreté ; il expose aussi les besoins de financement extérieur et les sources de financement connexes.

L'amélioration des modèles d'analyse, d'évaluation et de suivi de la pauvreté peut contribuer de manière décisive à l'élaboration des politiques appropriées et permettre aux décideurs de prendre des décisions reflétant la situation réelle des populations. De ce point de vue, il s'agit d'innover et de moderniser la façon dont les politiques sont élaborées et implémentées en termes d'actions, de projets ou de programmes. Ceci passe par une meilleure connaissance des caractéristiques des ménages, une définition appropriée des politiques à adopter vis-à-vis de chaque groupe cible, une évaluation et suivi des impacts des interventions publiques sur

les groupes cibles. Nous avons aussi besoin de mesures qui prennent en compte les informations pertinentes liées aux problèmes sociaux (tels que la santé public, l'éducation, le transport, le développement social, etc.) et la diversité humaine.

Les progrès vers l'éradication de la pauvreté et la réalisation des Objectifs du Millénaire pour le Développement (OMD) sont largement tributaires de l'amélioration des capacités des pays à intégrer, impliquer et inclure les personnes (et les évidences) au centre de l'élaboration et de l'implémentation des politiques de développement. Les défis majeurs des pays en développement ne peuvent être résolus simplement en allouant ressources financières car un pays (ou une région) peut disposer suffisamment de ressources financières et naturelles permettant à ses populations de vivre décemment et de s'épanouir dans la société, mais l'absence de politiques appropriées et légitimes peut considérablement entraver le développement des populations. Ainsi, nous avons besoin d'analyses approfondies des problèmes sociaux relatif à chaque société de manière à trouver une meilleure façon d'élaborer et de surveiller les politiques publiques. C'est dans ce but que nous avons préconisé dans cette thèse une méthodologie visant la construction des mesures multidimensionnelles significantes de la pauvreté (MDPMs).

Le mesurage de la pauvreté est un problème de décision dont on ne peut pas changer sa nature (c'est-à-dire l'ensemble des caractères, des propriétés qui la définissent), ni le séparer de son environnement. Ce problème est donc difficile à trancher selon une approche visant à trouver une *solution objective*, c'est-à-dire "qui est pauvre" et "qui ne l'est pas". La vérité est que le postulat reposant sur la recherche d'une *solution objective* (les "pauvres" et les "non-pauvres") ne peut en aucun cas être satisfaisant dans un problème où l'on souhaite concilier plusieurs attributs non commensurables et conflictuels, comme par exemple dans le cas d'une décision prenant en compte plusieurs aspects du bien-être. De plus, un tel postulat suscite d'autres questions difficiles telles que existe-t-il une ligne de démarcation absolue entre pauvre et non pauvre ? Comment définir le seuil de pauvreté sans tomber dans l'arbitraire ? Comment agréger l'ensemble des attributs pour décider qui est pauvre et qui ne l'est pas ? C'est fort de cela qu'une aide à la décision s'avère indispensable pour mieux appréhender les différentes facettes de la pauvreté et d'en proposer les politiques appropriées. L'aide à la décision sort du schéma de réflexion des approches classiques de mesurage de la pauvreté et propose des réponses au problème de pauvreté de part ses fondements méthodologiques. Elle foisonne des nuances qui permettent l'élaboration de nombreuses méthodes, voire des variantes de méthodes.

Le problème de mesurage de la pauvreté suscite trois conditions fondamentales : *la signifiante, l'opérationnalité et la légitimité*. La condition de signifiante concerne la manière dont on peut représenter certaines informations (notamment de nature qualitative) ayant trait à la pauvreté par un ensemble de valeurs numériques et cela de telle sorte que cette représentation reflète convenablement certaines pro-

priétés et axiomes de la pauvreté. La condition d'opérationnalité concerne la manière dont on peut agréger les informations ayant trait à la pauvreté par une valeur numérique unique et cela de telle sorte que cette représentation reflète convenablement le(s) politique(s) qui en découle(nt). La condition de légitimité concerne la manière dont on peut agréger les informations ayant trait à la pauvreté par une valeur numérique unique et cela de telle sorte que cette représentation reflète convenablement perception qu'ont les individus de leur situation réelle. C'est l'optique de construire des indices satisfaisant ces trois conditions fondamentales que nous avons développé dans cette thèse le concept de *mesure signifiante de la pauvreté*. Les conditions de *signifiante*, d'*opérationnalité* et de *légitimité* permettent de comprendre ce qu'implique l'opération de mesurage de la pauvreté et de connaître les traitements mathématiques autorisés et les outils d'aide à décision adéquats pour un niveau de mesurage donné. La méthodologie présentée dans cette thèse est basée sur deux grandes phases principales :

La première phase consiste à une classification non supervisée d'un échantillon de la population Ω dans le but de trouver une segmentation convenable et valide de la population en caractéristiques socio-économiques homogènes. Pour atteindre cet objectif, nous utilisons une technique statistique de classification (le clustering) pour la découverte des regroupements naturels des ménages lorsque l'on effectue des comparaisons quantitatives de leurs caractéristiques. Cette phase implique également la construction d'indices, le suivi-évaluation des capacités des ménages de chaque groupe et l'affectation des actions concrètes aux groupes identifiés. Ainsi, les groupes avec les politiques correspondantes sont le résultat des préférences exprimées par le(s) décideur(s) et reflètent la situation réelle des populations concernées. D'autre part, nous définissons aussi les conditions (tels que les jugements du client et des parties prenantes) sous lesquelles ce mesurage est possible. Les indices qui en découlent sont des mesures signifiantes, c'est-à-dire, théoriquement solide, opérationnellement complète conduisant à des résultats légitimes.

La deuxième phase consiste à une classification supervisée des ménages de l'univers U (la population totale) une ou plusieurs classes identifiées dans la première phase. Cette phase vise à formuler une procédure d'affectation de 'nouveaux' ménages de l'univers U à une ou plusieurs classes en examinant leurs caractéristiques et en se référant aux clusters diagnostics, aux indices d'admissibilité, aux indices d'éligibilité et aux indices de rejet que nous avons largement développé dans ce document. Nous supposons que les classes ne sont pas ordonnées. Ils sont plutôt décrits/caractérisés par une ou plusieurs 'modalités de référence' regroupées en sous-ensembles tel que chaque sous-ensemble, appelé '*cluster diagnostic*', soit associé à sa classe correspondante. De là nous affectons les nouveaux ménages aux classes spécifiques sur la base des clusters diagnostics. Pour atteindre cet objectif, nous définissons une nouvelle méthode de filtrage multi-attributs basée sur les principes de concordance et de discordance dérivés des concepts bien-connus de

majorité et de veto. Cette phase permet d'aboutir à des conclusions robustes, qui sont supportées par les données fournées et les axiomes, et sont en adéquation avec la théorie de mesurage.

Ce document comporte deux parties : l'état de l'art et la méthodologie générale de construction d'indices. L'état de l'art est composé de deux chapitres et la méthodologie générale de construction d'indices de deux chapitres. Nous présentons, ci-après, une brève description des cinq chapitres :

Chapitre 1. Préliminaire et problématique générale. Dans ce premier chapitre, nous présentons le contexte général du mesurage de la pauvreté. Puis, nous évoquons nos motivations, suivie de la présentation de la notion de mesure signifiante. Nous présentons enfin les objectifs généraux de la thèse.

Chapitre 2. Analyse et mesure de la pauvreté. Dans ce chapitre, nous présentons un peu d'histoire sur les origines, les fondements et la perception de la pauvreté et de son mesurage. Quelques questions liées au mesurage de la pauvreté y est évoqué ainsi qu'une présentation des approches existantes pour le mesurage multidimensionnelle de la pauvreté.

Chapitre 3. Méthodologie pour l'analyse et le mesurage de la pauvreté. Après une présentation des concepts de base, nous montrons comment transformer la table des réponses en matrice de proximité. Nous présentons ensuite notre procédure de classification et de construction des mesures signifiantes de pauvreté. Nous terminons ce chapitre par une illustration de l'implémentation des actions concrètes suivie d'une discussion et la présentation de quelques propriétés et axiomes.

Chapitre 4. Méthode de filtrage multi-attributs. Nous abordons d'abord la question de construction des clusters diagnostics. Ensuite, nous présentons la nouvelle méthode d'affectation floue et nous terminons par un exemple illustratif.

Après la conclusion générale, une application de notre nouvelles méthodes sur la base de données ASSL 2007 du Burkina Faso et une présentation des techniques de clustering existantes sont développées dans les Annexes **A** et **B** respectivement.

Préliminaire et problématique générale

1.1 Contexte général

1.1.1 La place de la pauvreté dans le monde

La question de la pauvreté occupe une place très importante dans les diverses sociétés du monde actuel. C'est l'un des thèmes centraux de nombreux débats théoriques, empiriques et institutionnels, aussi bien dans les pays développés que ceux en voie de développement. Les conséquences de la pauvreté diffèrent selon la région du monde que l'on traite, allant de l'exclusion sociale dans les pays riches jusqu'à la malnutrition, la vulnérabilité et la mort dans les pays les plus pauvres. La pauvreté est le plus grave et le plus répandu (il affecte plus de trois milliards d'êtres humains) de tous les maux dont souffre l'humanité.

Les problèmes liés à la pauvreté sont de vrais maux pour la communauté internationale, raison pour laquelle en septembre 2000, par la déclaration du millénaire adoptée aux nations unies lors de la plus grande réunion des Chefs d'État et de gouvernements jamais convoquée, plus de 150 chefs d'États des pays, riches et pauvres, se sont engagés à faire de leur mieux pour éradiquer ce fléau. En juillet 2005, le sommet du G8 tenu en Ecosse portait également sur la réduction de la pauvreté dans les pays en développement et s'était soldé à Gleneagles par une promesse de 25 milliards de dollars à l'horizon 2010 pour aider à la réduction de la pauvreté dans le monde. En 2008, Ban Ki-moon (secrétaire général des Nations Unies) annonçait que 16 milliards de dollars seraient levés d'ici 2015 pour renforcer la lutte contre la pauvreté dans le monde, notamment en Afrique. Puis, à l'issue de la réunion du G20 au sommet de Londres, les chefs d'État et de gouvernement des États membres du G20 se sont convenus d'autoriser le Fond Monétaire International (F.M.I.) à vendre une partie de ses réserves d'or pour financer l'aide aux pays les plus pauvres.

La pauvreté est une réalité mondiale et sa réduction est un objectif prioritaire des politiques publiques. Le suivi de cette pauvreté comporte deux composantes : l'une statique et descriptive pour identifier et décrire les différentes formes de pauvreté, l'autre dynamique et temporelle afin d'en comprendre et suivre les évolutions. Les questions relatives à la mesure de la pauvreté font encore de nos jours

coller beaucoup d'encre dans de nombreux axes de recherche théoriques, empiriques et institutionnels, aussi bien dans les pays développés que ceux en développement. En effet, la diversité des agents vecteurs de la pauvreté et la complexité de ses manifestations opposent une résistance logique aux différentes stratégies de lutte contre la pauvreté adoptées.

Le concept de pauvreté peut être expliqué en sociologie, mais les outils mathématiques classiques utilisés ne peuvent pas appréhender la portée réelle du phénomène. En effet, l'aspect multidimensionnel de la pauvreté, son caractère flou et les substitutions possibles entre ses différentes formes, rendent la recherche et le suivi des facteurs explicatifs assez complexes. Selon Sen [2001], *“la pauvreté est un monde complexe, multiforme, qui exige une analyse précise de toutes ses nombreuses dimensions”*. Il est alors indiqué d'entreprendre une démarche opérationnelle qui ait pour vocation de prendre en compte l'ensemble de ces aspects pour mieux se référer à la réalité. Certaines réactions face à ce fléau mettent un accent prioritaire sur le soutien à accorder aux pauvres ; d'autres, par contre, dénoncent le mépris des pauvres et les inégalités sociales. Toutefois, il n'en reste pas moins qu'il est nécessaire, avant de mener toute action de réduction de la pauvreté, d'en offrir une définition opérationnelle et une méthode de mesure qui permettent d'appréhender le phénomène.

Malgré l'existence de plusieurs mesures conçues pour capturer la pauvreté aiguë des personnes, telles que le revenu par habitant, PIB par habitant, l'indice HDI ou l'indice MPI, la pauvreté continue à sévir dans la plupart des régions du monde et ses conséquences sont incalculables. On observe qu'une majorité des pays les plus pauvres des mondes sont aujourd'hui en Afrique, en dépit d'une richesse en ressources naturelles. Certes, certains pays africains comme l'Afrique du Sud et l'Égypte ne sont pas aussi pauvres que d'autres comme le Niger et l'Éthiopie. En 2010, 22 des 24 nations identifiées comme ayant un “Faible Développement Humain”, au regard de l'indice de développement humain développé par le PNUD, ont été localisées en Afrique sub-saharienne. D'autre part, la plupart des pays présentant les taux les plus élevés de pauvreté selon l'indice MPI la même année se trouve en Afrique. En 2006, 34 des 50 nations sur la liste de l'ONU des pays développés les moins avancés (LDCs¹) se trouvent en Afrique. Il est vérifié que la pauvreté multidimensionnelle aiguë (selon l'indice MPI) est plus prononcée dans les pays africains que d'autres pays du continent. Ce qui est plus inquiétant, c'est que la situation ne semble pas s'améliorer malgré les programmes d'aide au développement lancés dans les pays concernés.

Cependant, les travaux de Alkire et Santos [2010] montrent que la pauvreté multidimensionnelle n'est pas seulement un problème africain, mais elle affecte

1. LDC : Least Developed Countries, voir <http://www.un.org/special-rep/ohrlls/lcd/statistics.htm>

aussi les autres pays du monde entier. Par exemple, l'Asie du Sud présente les plus hauts niveaux de pauvreté du monde en terme de niveau de vies humaines. Les pays d'Amérique latine, des Caraïbes et de l'Asie orientale et du Pacifique révèlent également une proportion élevée de pauvres selon l'indice MPI. Un projet de recherche à long terme achevé au sein de l'Observatoire Européen sur la Situation Sociale [voir [Lelkes et Zólyomi, 2008](#)] mentionne que le taux de pauvreté varie entre 10 % et 23% dans les pays de l'Union Européenne. Les faibles niveaux de pauvreté caractérisent les pays "Scandinaves", les pays "Corporatistes" (tels que Autriche, Allemagne), et la République Tchèque, la Slovaquie et la Slovénie parmi les pays de l'ex-socialistes. En revanche, le risque de pauvreté tend à être relativement élevé dans la région méditerranéenne et les pays baltes. Au total, environ 75 millions de personnes dans l'Union Européenne sont à risque de pauvreté. Les pays ayant la plus forte population pauvre sont entre autres France, l'Allemagne, l'Italie, la Pologne, l'Espagne et le Royaume-Uni. Selon [Smith \[2010\]](#), le U.S. Census Bureau, approximativement 43.6 millions d'euros (14.3%) d'américains vivaient dans la pauvreté absolue en 2009, au dessus de 39.8 millions (13.2%) en 2008.

Ces résultats montrent que la pauvreté demeure un fléau du monde entier malgré la variété des mesures et approches existantes dans la littérature. Pourquoi cet échec 'catastrophique' pour éradiquer la pauvreté ? Pourtant, la pauvreté est connu de tous ! Nous pensons qu'il s'agit entre autre d'un

- **problème de méthodologique** : Quelle approche méthodologie adoptée pour l'élaboration et l'implémentation des politiques publiques appropriées ?
- **problème de volonté politique** : la preuve apparaît au regard de la répartition actuelle des richesses fonctionnelles pour maintenir le statu quo géopolitique mondiale. Les gouvernements sont-ils VRAIMENT engagés dans la poursuite de politiques adéquates de réduction de la pauvreté ?
- **problème de volonté sociale** : les sociétés les plus riches sont-ils conscients et prêts à transférer une partie de leur richesse aux sociétés les moins riches ?
- **problème culturel** : Sommes-nous sûrs que la réduction de la pauvreté implique augmentation de la consommation (de biens et services) en utilisant les normes de consommation des sociétés les plus riches ?

C'est au regard des problèmes sus-cités que nous avons introduit le concept de "mesures significantes" que nous allons développer dans cette thèse. L'idée consiste à proposer une méthodologie générale concernant la façon dont la science peut aider dans la lutte contre la pauvreté et dans la conception et l'implémentation des politiques adéquates de réduction de la pauvreté.

L'examen de la littérature nous montre que la plupart des politiques actuelles de réduction de la pauvreté sont basées sur des caractérisations simplistes de qui est *pauvre* et qui ne l'est pas. Autrement dit, la plupart des approches de mesure de la pauvreté ne visent pas à aider à concevoir et implémenter des politiques, mais à la représentation d'une "situation objective de la pauvreté". Malheureusement,

nous avons vu que cela est trompeur. Mesurer la croissance du PIB d'un pays ne permet pas de comprendre si ses citoyens sont de moins en moins pauvres. Les informations qualitatives sur les conditions de vie de la population sont souvent négligées par les indices de pauvreté surtout lorsqu'elles ne peuvent pas être quantifiées. L'importance d'offrir des chances à chaque individu d'améliorer ses opportunités et ses capacités n'a été sérieusement envisagé que par la théorie des capacités de Sen. Mais, cette approche a été très souvent accusée d'être difficilement opérationnalisable. Nous envisageons également dans cette thèse de montrer comment opérationnaliser l'approche par les capacités au travers des mesures signifiantes. Notre observation est qu'au-delà des raisons politiques, sociaux et culturels pour lesquels nous ne parvenons pas à être efficaces dans la réduction de la pauvreté, nous manquons aussi des outils de mesure adéquat, capables d'aider à la conception et l'implémentation des politiques de réduction de la pauvreté. Nous manquons d'outils d'analyse des politiques dans ce domaine critique.

Avant d'aborder une discussion plus large sur la question de savoir comment nous pouvons améliorer la mesure multidimensionnelle de la pauvreté grâce aux mesures signifiantes, nous présenterons quelques grandes caractéristiques tirées de la littérature montrant que la pauvreté est un *problème complexe*.

1.1.2 La complexité du problème

1.1.2.1 La pauvreté est multidimensionnelle

Plusieurs auteurs [tels que [Townsend, 1987](#); [Sen, 1979a, 1985, 1993, 2001](#); [Bourguignon et Chakravarty, 2003](#); [Fusco, 2005](#); [Bertin, 2007](#); [Kakwani et Silber, 2008](#); [Sami et Lahga, 2008](#); [Alkire et Santos, 2010](#); [Alkire et Foster, 2011](#)] et organisations internationales [[World Bank, 2001, 2005](#); [UNDP, 2010](#)] s'accordent qu'un seul indicateur ne peut pas capturer les multiples facettes que constituent la pauvreté. La pauvreté est multidimensionnelle. Elle comprend plusieurs facteurs tels que la santé, le manque d'éducation, le niveau de vie inadéquates, le manque de revenu, le manque de représentation et de la liberté, le sentiment d'impuissance, la mauvaise condition de travail et la violence. Ainsi, lorsque nous construisons une mesure de la pauvreté, nous devons d'abord accepter sa nature multidimensionnelle. Deuxièmement, nous devons développer une mesure multidimensionnelle de la pauvreté capable de soutenir les processus d'élaboration des politiques : elle doit contribuer à l'élaboration des politiques appropriées en basant sur des informations disponibles. Cependant, les politiques peuvent être également multidimensionnelles : ils peuvent cibler différents groupes dans le même temps, se concentrer sur la réalisation des objectifs multiples, prendre en compte des priorités multiples. Une politique de réduction de la pauvreté ne se limite uniquement aux efforts visant à augmenter le revenu de la population. Sous une telle perspective, la mesure

de la pauvreté doit être un instrument pour l'élaboration et l'implémentation des politiques, et elle doit permettre d'évaluer l'efficacité des politiques actuelles et de déterminer si la situation évolue. Le MPI proposé par [Alkire et Santos \[2010\]](#) a été développé avec cette idée à l'esprit. Le problème consiste alors à construire une mesure multidimensionnelle de la pauvreté (MPM), qui pourrait intégrer une série d'indicateurs capturant la situation et permettant d'élaborer des décisions éclairées au sujet des politiques, programmes ou projets visant à éradiquer la pauvreté.

Du point de vue technique, le problème auquel fait face toutes les tentatives de construire un indice de pauvreté multidimensionnelle est de savoir comment agréger les informations disponibles portant sur plusieurs attributs (accès aux services de santé, le revenu, l'eau, l'assainissement, l'éducation, nourriture, etc) en un indice composite unique. D'autre part ceci doit être fait de façon significative [voir [Roberts, 1979](#)], d'autre part ça doit être fait de telle manière que l'indice puisse vraiment être utilisé afin de soutenir la conception, l'implémentation et l'évaluation des politiques de réduction de la pauvreté. Bien que certains indices (tels que MPI et HPI) proposés par la littérature évoque la nécessité de la prise en compte de la signifiante dans la conception de la procédure d'agrégation, il est difficile de savoir si ces indices sont vraiment capable de remplir la condition d'opérationnalité. La plupart du temps, ceci est dû à l'idée trompeuse selon laquelle un indice de la pauvreté à usage général pourrait être utilisé pour l'élaboration et l'implémentation de tous types de politiques de réduction de la pauvreté.

1.1.2.2 La pauvreté est évolutive

Ce qui nous permet de connaître la situation des personnes - durant le processus de mesure de la pauvreté - varie dans le temps selon les sociétés. En outre, les manifestations de la pauvreté sont intimement liées aux structures de la société et la période pendant laquelle la pauvreté est discutée. De ce point de vue, un pauvre dans un pays développé (comme les États-Unis ou la France) ne doit pas être considérée de la même manière que les pauvres dans les pays en développement (comme l'Asie du Sud ou l'Afrique sub-saharienne) ou les pays les plus pauvres d'Amérique latine. Une mesure multidimensionnelle de la pauvreté doit exécutée dans une période donnée dans un espace bien défini et dans un contexte particulier. Ainsi, les modèles théoriques et empiriques, basés sur la capacité des ménages à satisfaire les besoins de base doivent incorporer les faits ou évidences provenant de l'analyse sociale.

Sous une telle perspective, il est clair que les politiques réduction de la pauvreté sont contingentées. Les politiques conçue pour être appliqué, disons par exemple en Afrique, ne sont pas nécessairement applicable, disons par exemple en Amérique latine et sûrement pas dans la même période. Connaître la pauvreté n'est pas uniquement une question de mesurer un phénomène quantifiable (combien de mé-

nages n'ont pas accès à l'eau potable ou agrégées les données au niveau pays), mais aussi une question de tendance et d'ampleur sociale de ce que nous observons. La conception des politiques nécessite une telle connaissance autant qu'elle a besoin d'informations quantifiées. Cet aspect est rarement pris en compte par les indices de pauvreté existants.

1.1.2.3 La pauvreté est une situation non objectivable

Outre l'information objective (comme le revenu, l'accès aux services de base ou la qualité du logement) qui pourrait caractériser les ménages "pauvres", il y a une dimension subjective de la pauvreté. Un ménage sans robinet à eau potable dans la maison se sentira pauvre, si tous les autres ménages dans le village disposent d'un robinet à eau potable dans leur maison. En effet, si personne dans le village n'a accès à une telle installation alors ce manque de service ne sera pas considérée par les ménages comme une "pauvreté" même si objectivement parlant (et dans le respect des standards normaux de richesse), c'est un signe de pauvreté. Le même raisonnement s'applique pour de nombreux autres types de situations observables, notamment les symptômes inexplicables de sentiment d'être pauvre : par exemple, ne pas posséder d'une radio en Afrique est un signe de l'extrême pauvreté...

Les individus ne sont pas seulement pauvres, ils se sentent également plus ou moins pauvres comparativement au contexte économique, la culture, l'histoire et les traditions locales. De plus, les ambitions de chaque individu sur la façon de sortir de la pauvreté sont différentes d'un cas à l'autre. Certains peuvent aspirer pour une meilleure éducation pour leurs enfants, tandis que d'autres à un meilleur accès à l'eau potable et aux installations sanitaires ou aux services de santé, tous ces aspirations étant légitimes (mais pas nécessairement réalisables). Cette dimension subjective de la pauvreté a été capturée par Sen dans sa théorie des capacités : un individu *est* (ou *se sent*) "pauvre" parce qu'il n'est pas en mesure de réaliser ses aspirations, ces derniers étant subjectivement définis. En concevant une politique de réduction de la pauvreté, il est donc nécessaire d'être en mesure de prendre en compte cette dimension.

1.1.2.4 Différentes pauvretés

La discussion précédente nous amène à faire une dernière observation. Il est trompeur de ne parler que de "pauvres" et de "non-pauvres", au moins lorsqu'une perspective multidimensionnelle de la pauvreté est considérée. Ce que nous observons dans la réalité sont les différents types de pauvreté. Nous observons les individus en situation de manque plus ou moins étendue, avec différentes sévérités et différemment perçue d'accès aux biens et services, d'opportunités ou de sécurité. Certes, il y a des individus n'ayant pas accès à (presque) tout (l'extrême pauvreté),

mais encore une fois cela peut dépendre des pays, des cultures et des traditions. En outre, la situation d'extrême pauvreté est beaucoup moins importante dans la population qui est intuitivement pauvre, mais pas encore très pauvre. Au regard de la perspective d'élaboration des politiques, on remarque qu'il existe une vaste majorité de plus ou moins pauvres, lesquels doivent être considérés comme la cible des politiques.

Il est objectivement difficile de construire une seule catégorie de pauvres. De plus, peut-être cela n'a aucun sens de le faire. Après tout, pourquoi avons-nous besoin de classer certaines personnes aussi pauvres ? Positivement parlant la raison en est que ces "personnes pauvres" vont être la cible d'une certaine politique visant à les aider à sortir de cette situation. Mais alors, pourquoi avons-nous besoin d'une seule catégorie de pauvres ? Peut-être qu'il est plus sensé d'avoir différentes catégories de personnes différemment pauvres à partir desquels on peut construire les politiques adéquates. Car il est vraisemblable d'avoir des politiques plus efficaces si elles sont bien ciblées.

1.2 La pauvreté et son mesurage

1.2.1 Qu'est ce que la pauvreté ?

Lutter contre la pauvreté implique de mesurer le phénomène et d'en déterminer les causes les plus profondes pour proposer des politiques efficaces. Aussi, pour bien mener une politique de lutte contre la pauvreté, il faut bien identifier les "pauvres", et cette identification est tributaire de la définition qu'on donne au concept de pauvreté et partant, du type de mesure utilisée. La première difficulté de la recherche dans le domaine de la pauvreté est de trouver une définition universelle ou normalisée de ce phénomène car il n'existe pas une définition de la pauvreté faisant l'unanimité. Plusieurs auteurs ont suggéré différentes définitions depuis la genèse de ce fléau jusqu'à nos jours. Mais, ces définitions sont restées tributaires du contexte, de la période et surtout, de la compréhension qu'ils se font de la pauvreté. C'est ainsi que Encarta World English Dictionary [Encarta, 2010] définit la pauvreté comme "l'absence de besoins humains fondamentaux, tels que l'eau potable, la nutrition, l'accès aux services de santé, l'éducation, les vêtements et le logement, en raison de l'incapacité de moyens. Le P.N.U.D. (Programme des Nations Unies pour le Développement) la considère comme "*la négation des opportunités et des possibilités de choix les plus essentielles du développement humain*". Watts [1968] la considère comme "*l'absence de maîtrise des marchandises de base, en général*". De manière plus générale, Sen [1985] la définit comme le manque de 'capabilité' à fonctionner dans une société donnée et la banque mondiale [World Bank, 2001] comme une déprivation prononcée en bien-être. Toutes les

définitions montrent que les auteurs s'accordent sur le fait que le *pauvre* est celui à qui il manque "*quelque chose*" considérée comme utile pour atteindre un niveau de vie socialement acceptable. Cette "*chose*" manquante se différencie du *niveau d'un indicateur* (pays, région, communauté, ménage ou individu) à l'autre dans une période donnée, et se décrit de différentes manières. Ainsi, une personne qui s'en tire bien selon les normes d'un pays en développement pourrait très bien être considérée comme désespérément pauvre dans un pays développé. De même, une personne avec une grande villa, sans véhicule pourra se dire pauvre tandis qu'un autre vivant dans une case, sans mobilier ni voiture pourra se dire comblée.

L'échec lié à une tentative de définition du concept de pauvreté est une réalité. On pourrait être tenté de définir comme *pauvre*, toute personne se trouvant dans l'incapacité de satisfaire les besoins fondamentaux d'un être humain. Mais, là encore, un *besoin* dépend de la vision que des gens possèdent du monde. Parfois nous sommes si prompts à qualifier une personne de "pauvre" selon notre vision du monde (ou selon la vision universelle), plutôt que selon la sienne. En guise d'exemple, dire qu'un pygmée de la forêt équatoriale du Sud-Cameroun est pauvre simplement parce qu'il est couvert d'une écorce d'arbre comme cache-sexe ou qu'il habite une hutte construite à base de feuilles d'arbre ; parce qu'il ne consomme pas une eau potable ou n'a pas un moyen d'assainissement adéquat dans son village ; parce que son village est dépourvu d'une école ou d'un centre de santé, serait une erreur monumentale. La "*vérité*" est que ce pygmée a sa manière de concevoir les choses, entre autres le concept de pauvreté. Pourtant, un pygmée, vivant dans sa forêt, se trouve dans un paradis, ne se souciant de rien : telle est, en réalité, sa vision du monde, très différente de la nôtre ! Le priver de la forêt serait un crime et accélérerait sa mort ! Il s'agit donc de la *pauvreté subjective* qui est l'état d'une personne ne disposant pas de ce dont il a besoin (que ce soit matériel ou culturel) et d'un environnement adéquat, selon lui, pour son épanouissement.

Le terme "*pauvre*" indique forcément une place dans la société qui est identifiable par des critères matériels et sociaux. Cette place occupée par un individu dans la société n'est pas volontaire, encore moins accidentelle ; mais, elle est déterminée par la structure socio-économique. Autrement dit, pour bien saisir le concept de pauvreté, nous devons projeter les individus dans l'espace des critères socio-économiques à partir des relations qu'ils entretiennent avec leur environnement. Avec la mondialisation et les progrès scientifiques, huit pays se sont rangés sous l'acronyme *pays du G8* où la pauvreté est caractérisée par l'exclusion sociale ; d'autres pays, par contre, se sont hissés au rang de *pays en voie de développement* avec une pauvreté en voie de disparition ; et, les pays sous-développés appartenant à une classe de pays appelée *pays du tiers-monde*, où la pauvreté est caractérisée par la *pauvreté matérielle* et la *pauvreté des potentialités*.

Remarque 1.2.1 *La pauvreté matérielle (pécuniaire) se caractérise par le manque*

de services d'eau, d'assainissement et d'hygiène, la malnutrition, les épidémies, l'absence de services sanitaires et d'assurance médicale, l'absence des systèmes éducatifs appropriés. Tandis que la pauvreté des potentialités exprime le manque de moyens permettant de sortir de la pauvreté. La pauvreté matérielle et la pauvreté des potentialités mettent l'accent sur ce qui manque et les conséquences qui en découlent de cet état de "manquement".

Parce que les êtres humains sont complètement différents et subjectifs, la pauvreté doit être examinée à travers une variété d'indicateurs. C'est pourquoi Sen [2001] a lors d'un interview déclaré que : *"la pauvreté est un monde complexe, multiforme, qui exige une analyse précise de toutes ses nombreuses dimensions"*. Allant dans le même sens, les nations unies et la banque mondiale [World Bank, 2001, 2005] ont proposé deux définitions alternatives que nous avons résumé comme suite :

(i) *"poverty is a denial of choices and opportunities, a violation of human dignity. It means lack of basic capacity to participate effectively in society. It means not having enough to feed and cloth a family, not having a school or clinic to go to, not having the land on which to grow one's food or a job to earn one's living, not having access to credit. It means insecurity, powerlessness and exclusion of individuals, households and communities. It means susceptibility to violence, and it often implies living on marginal or fragile environments, without access to clean water or sanitation"* (UN Statement, June 1998 - signed by the heads of all UN agencies).

(ii) *"la pauvreté : c'est perdre un enfant suite à une maladie survenue par la consommation d'une eau non potable. C'est ne pas pouvoir bénéficier d'une éducation et être vulnérable (manque d'argent et de ressources). C'est la famine, le manque d'abri, le manque d'opportunité d'emploi ou de biens physiques tels que le bétail et la terre ; c'est être malade et ne pas être à mesure d'aller en consultation ou d'acheter ses médicaments ; c'est être sans emploi, avoir le sentiment d'impuissance. C'est la peur du futur, 'vivre au jour le jour', c'est le manque de voix, de représentation ou de liberté"* [World Bank, 2001, 2005].

Les définitions (i) et (ii) prennent en compte plusieurs indicateurs (tels que le niveau de revenu et de consommation, les indicateurs sociaux, les indicateurs de vulnérabilité aux risques) et les indicateurs d'accès à la vie sociopolitique) et semblent plus représentatives de la situation du bien-être humain qu'une définition purement monétaire (définition unidimensionnelle). La pauvreté a plusieurs facettes, elle doit donc être définie et mesurée de manière multidimensionnelle.

1.2.2 Pourquoi mesurer la pauvreté ?

La mesure de la pauvreté semble inévitable à cause de son importance dans plusieurs contextes, tels que la lutte contre la pauvreté et l'exclusion sociale, la promotion d'égalité sociale, la défense des droits des non-privilegiés, la compréhension de la vulnérabilité, la lutte contre les inégalités. Dans cette thèse, nous avons identifié au moins 5 raisons principales² [voir [Haughton et Khandker, 2009](#); [World Bank, 2005](#)] pouvant justifier la nécessité d'un mesurage rigoureux de la pauvreté :

Amélioration de notre connaissance de la situation

[Ravallion \[1998\]](#) a écrit que *“une mesure crédible de la pauvreté peut être un instrument puissant pour focaliser l'attention des décideurs sur les conditions de vie des pauvres”*. La première raison du mesurage de la pauvreté est l'identification des pauvres. Le mesurage de la pauvreté devrait permettre d'augmenter notre connaissance des pauvres, comment la pauvreté a été perçue au cours du cycle de vie et comment elle se reproduit. La recherche des causes et des caractéristiques de la pauvreté dans un pays ou une région donnée est donc une étape analytique cruciale pour l'élaboration des politiques de réduction de la pauvreté et la mise sur pied des stratégies de lutte. En d'autres termes, le mesurage de la pauvreté rend visible les nécessités grâce à l'examen des facteurs déterminants de leurs conditions de vie. Il permet ainsi des comparaisons de niveaux de pauvreté dans le temps et dans l'espace.

Ciblage des interventions nationales et internationales

Après l'identification des “pauvres”, le mesurage de la pauvreté joue également un rôle capital dans la conception et le ciblage des interventions de lutte les mieux adaptées à la situation. Généralement, on utilise des profils de pauvreté [voir [Ravallion, 2008](#); [Chen et Ravallion, 2007](#)] pour atteindre cet objectif. L'idée consiste à examiner comment une mesure de la pauvreté varie selon les sous-groupes d'une population (par exemple, par village, ville, région, etc.) et de comparer les principales caractéristiques des pauvres par rapport aux non-pauvres.

Développement d'une stratégie de dépenses sociales et la croissance

Le mesurage de la pauvreté permet d'évaluer si la croissance économique générale a eu une influence positive chez les pauvres, c'est-à-dire, si elle a contribué à l'amélioration de leurs conditions de vie. Généralement, on utilise l'information sur les ménages et leurs status économiques afin de comprendre les effets positifs

2. Notons que la liste de ces raisons ne sont pas exhaustives, il s'agit juste des plus importantes.

ou négatifs de l'ensemble des politiques économiques et financières implémentées en actions concrètes sur le terrain. Ainsi, selon que la pauvreté est croissante ou décroissante, les décideurs pourraient ajuster les réformes et réévaluer comment les pauvres sont affectés par ces réformes. Par exemple, nous pouvons changer la politique fiscale et réévaluer ensuite comment ces changements influencent la conditions de vie des populations pauvres.

Suivi et évaluation des projets et des interventions politiques

La quatrième raison concerne l'évaluation de l'effectivité des politiques (programmes or projets) actuels d'aide aux personnes pauvres. De ce fait, le mesurage de la pauvreté est perçu comme un instrument de jugement des retombés (effets positifs vs effets négatifs) d'une politique d'amélioration des conditions de vie des populations pauvres. Il permet alors d'évaluer les résultats obtenus en comparaison avec un groupe de référence. Le mesurage de la pauvreté joue donc un rôle important dans l'amélioration des politiques mise en œuvre par des actions concrètes et la mise à jour des mécanismes de conception des projets ou des programmes. La simulation de l'impact des politiques alternatives sur la pauvreté et la sélection des mécanismes appropriés des différentes politiques sont entre autre les deux principales raisons de cette étape.

Evaluation de l'effectivité des institutions

Le mesurage de la pauvreté peut aussi être perçu comme un outil d'évaluation des institutions de lutte contre la pauvreté. Il s'agit notamment de savoir si, oui ou non, une institution ou un gouvernement accomplit correctement la tâche qui lui a été assignée, c'est-à-dire la réduction de la pauvreté. L'idée consiste alors à évaluer le succès d'une institution ou d'un gouvernement dans la poursuite de la lutte contre la pauvreté. Le mesurage de la pauvreté est alors utile dans la mesure où il permet d'évaluer combien la pauvreté a décru et la durabilité des résultats dans le temps et dans l'espace.

1.2.3 Vers une nouvelle méthodologie de mesurage

Lutter contre la pauvreté implique de mesurer le phénomène et d'en déterminer les causes les plus profondes pour proposer des politiques efficaces. Aussi, pour bien mener une politique de lutte contre la pauvreté, il faut identifier les différentes classes de pauvreté, et cette identification est tributaire de la définition qu'on donne au concept de pauvreté et partant, du type de mesure utilisé. Certes, les seuils monétaires sont incontestablement utile pour cerner une dimension importante du bien-être, mais ne permettent pas de rendre compte de toute la complexité qui en-

ture la pauvreté. La justification “métrique” ou “quantitative” de l’état des privations basée sur l’approche monétaire ne fait que masquer la difficulté de construire une mesure objective. En tenant compte de la diversité des individus, Sen [1992] considère que l’évaluation de la pauvreté sur la base des ressources ou des dotations est inadaptée pour juger l’état des privations. C’est dans cette optique qu’il souligne que : “*nous nous distinguons les uns des autres par des caractéristiques non seulement externes (la fortune reçue en héritage, le milieu naturel et social où nous vivons), mais aussi personnelles (l’âge, le sexe, la vulnérabilité aux maladies, les aptitudes physiques et intellectuelles)*”.

De ce point de vue, nous devons prendre en compte la diversité humaine dans le processus de mesure de la pauvreté. Omettre un tel aspect peut conduire à une remise en question de la validité du modèle et l’inadéquation des politiques de décision avec les résultats attendus. Ainsi, deux individus également dotés en fortune personnelle ne sont pas forcément sujet à une même politique car ils peuvent ne pas disposer des mêmes conditions de vie, en termes de libertés, de logement ou de santé. Ceci montre l’importance de travailler dans un espace plus large et multidimensionnel : celui des capacités. L’idée des capacités de Sen nous amène à considérer la pauvreté comme “*un manque fondamental de capacités minimales adéquates*” empêchant l’individu de réaliser le type de vie qu’il souhaite et l’entraînant dans l’indigence. L’analyse du bien-être doit donc prendre en compte cette remarque pour établir une mesure exhaustive de ce phénomène, tant du point de vue des causes les plus profondes que des politiques de lutte contre la pauvreté.

Certains auteurs [à l’instar de Sen, 1985, 1992, 1993, 1999; Sami et Lahga, 2006; Fusco, 2005; Bourguignon et Chakravarty, 2003; Bertin, 2007; Alkire et Santos, 2010; Alkire et Foster, 2011] se sont démarqués des approches traditionnelles (en termes purement monétaires et de la consommation) pour embrasser et encourager l’analyse du bien-être par les approches multidimensionnelles. Bien que l’apport théorique des approches multidimensionnelles soit remarquable et fasse aujourd’hui consensus, le problème de leurs opérationnalisations suscite encore de nombreux débats et demeure de nos jours un défi à relever. De nombreuses questions se posent lorsqu’on entame le processus de modélisation de la pauvreté : combien de classes de pauvreté existe-il ? Comment décrire ces classes et y associer des politiques adéquates ? De quelle manière faut-il évaluer le bien-être des citoyens pour les affecter aux classes ? Comment évaluer le bien-être d’un groupe d’individus pour décider d’une politique de réduction de la pauvreté adaptée au groupe cible. Une enquête auprès des ménages fournit-elle toujours des indications fiables ? Quelle mesure doit être utilisée pour agréger les données sur le bien-être des citoyens ? Le choix de cette mesure importe-t-il ? Est-il possible de déterminer le coût minimum nécessaire pour éradiquer/supprimer la pauvreté dans une communauté spécifique ? Quelles politiques d’action adopter pour un sous-groupe donné ? Le choix de cette politique importe-t-il ? Ces questions posent, de façon gé-

nérale, les problèmes d'analyse des données et du choix de la méthodologie d'aide multicritère à la décision. C'est en vue de combler les insuffisances mentionnées ci-haut que cette thèse a été initiée.

L'évaluation (ou mesure) de la pauvreté est une suite ordonnée d'opérations ou de pratiques très diverses organisées dans le but d'effectuer un jugement de valeur, en un moment donné, sur l'état de pauvreté du niveau d'un indicateur. C'est la confrontation d'un résultat à un objectif. La mesure de la pauvreté est donc un processus de recueil, puis de traitement de données pertinentes sur la pauvreté du niveau d'un indicateur, qui ne prennent sens qu'au regard de la décision qui fonde l'utilité de la mesure. Le résultat d'une telle mesure doit toujours être la réponse à un objectif spécifique de réduction de la pauvreté et doit aussi prendre en compte la subjectivité des acteurs qui la commandite. La mesure de la pauvreté peut donc servir différents objectifs et différentes approches de lutte contre la pauvreté.

Dans le cadre de *l'approche analytique*, l'objectif d'une évaluation de la pauvreté est d'identifier et d'étudier les facteurs de variations de ce fléau. En effet, les manifestations de la pauvreté et le bien-être sont étroitement liés aux structures de la société dans laquelle elle est étudiée. Aussi, la question de pauvreté et ses manifestations dépendent du lieu et l'époque où elle est considérée. Dès lors, une approche analytique peut-être utilisée aux fins de recherche, comme par exemple pour décrire l'impact de la privation physiologique ou sociologique sur l'état des personnes et leur comportement. L'objectif de ce type d'approche analytique est de mieux cerner la dynamique de pauvreté afin d'appréhender les besoins des personnes, leurs préférences et leurs aversions. De tels travaux contribuent, d'une part, à la compréhension du phénomène de pauvreté dans la société étudiée selon une époque donnée. D'autre part, ces travaux peuvent alors aboutir à une augmentation des connaissances, puis à des solutions pratiques (et plus réaliste) permettant de résoudre certains problèmes de bien-être.

Dans le cadre d'une logique d'intervention et d'aide à la décision, l'objectif est d'avoir une vue d'ensemble du phénomène de pauvreté à travers des observations exhaustives. En effet, du fait que la pauvreté soit difficilement objectivable, ayant une nature auquel il est difficile de donner des frontières, des observations exhaustives contribueront à identifier d'éventuels problèmes, puis leurs causes, pour pouvoir proposer des mesures correctrices appropriées. Enfin, la mesure de la pauvreté peut être réalisée dans le cadre de la conception des systèmes, afin par exemple de choisir le système innovant maximisant un certain nombre d'indicateurs de la pauvreté sur lesquels différents systèmes "*candidats*" auront été testés.

1.3 Nos motivations

1.3.1 Les besoins d'opérationnalisation des capacités

La revue de la littérature [cf. Kana *et al.*, 2011] sur la mesure de la pauvreté montre que le mesurage de la pauvreté n'est pas une représentation d'une situation objective, mais plutôt un instrument de poursuite d'une certaine politique. En effet, on peut se sentir 'pauvre' et ne pas être identifié comme tel, ou alors on peut être identifié comme 'pauvre' et ne pas se sentir comme tel. On observe en réalité que les individus sont plus ou moins pauvres et de différentes manières sur dimensions différentes. La pauvreté est un concept évolutif, multidimensionnel, flou et non objectivable qui ne contient rien de numérique, mais seulement la sensation de ceux qui en souffrent.

Malgré les nombreuses critiques émises contre les approches traditionnelles d'analyse et de mesure de la pauvreté, plusieurs auteurs [see Nussbaum, 1987, 2000; Fusco, 2005; Alkire, 2005; Bertin, 2007; Kana *et al.*, 2011] s'accorde que l'approche par les capacités [see Sen, 1985] est plus appropriée pour évaluer le bien-être des individus principalement parce qu'elle met en exergue la diversité des relations entre les individus et les *marchandises*³ ('*paniers de biens et de services*'), la complexité des relations qu'entretiennent les individus entre eux (relations sociales) et les individus avec leur environnement (institutions, normes, cultures). L'argument *fort* de l'approche par les capacités repose sur le postulat stipulant que les marchandises (ou les ressources) des individus ne peuvent suffire pour décrire fidèlement leur bien-être. En guise d'exemple, deux personnes peuvent aspirer aux *choses* différentes, en termes de bien-être, avec les mêmes ressources équivalentes à, supposons 3500 €.

Sen [1985] se contente alors d'utiliser les "*caractéristiques*" des marchandises plutôt que les marchandises (ou les ressources) des individus en fondant son analyse sur la *valeur* des paniers de biens (ce que *sont* ou *font* les individus grâce aux caractéristiques des biens) plutôt que sur la référence usuelle des *biens*. C'est pourquoi Sen [1985] a fait une distinction claire entre les *accomplissements* d'une personne (*person's interests*) et ses *opportunités*, respectivement appelés "bien-être" et "avantage". Sen mentionne que "*well-being is concerned with a person's achievement : how 'well' is his or her 'being' ? 'Advantage' refers to the real opportunities that the person has, especially compared with others*". Ce postulat considère donc les ressources (biens et services) des individus comme le moyen d'améliorer leurs qualités de vie et préconise de se focaliser sur la manière dont ces individus vont utiliser leurs ressources (qui sont, par hypothèse, équivalentes). Cette idée de Sen conduit au développement d'une discussion plus large concernant les distinctions entre *commodités*, *caractéristiques*, *fonctionnements* et *capacités*. L'ap-

3. En anglais, "commodities".

proche de Sen permet ainsi de prendre en compte la notion de *liberté* d'un individu à choisir d'atteindre un niveau élevé de bien-être, et de considérer l'hypothèse de la *diversité humaine* durant le processus de mesurage de la pauvreté.

Cependant, l'approche par les capacités de Sen a suscité de nombreuses critiques dans la littérature [cf. Navarro, 2000; Pogge, 2002; Nussbaum, 1987] dont les plus virulentes sont ceux de Bénicourt [2004, 2006]. Les critiques portent généralement sur les ambiguïtés de la théorie des capacités et la mise en doute de la possibilité d'agrèger en un indice unique les multiples capacités. La question qui suscite les débats est : *comment évaluer les capacités des individus ?* Certes, une telle évaluation, si elle est entreprise, doit prendre en compte différentes dimensions d'incertitude. Nous devons *choisir* et *valider l'espace des fonctionnements* que les individus sont capables de "*faire*" (doing) ou aspirent à "*être*" (being) grâce à leurs ressources et à leurs caractéristiques. Le *choix* et la *validation* de l'espace de *fonctionnements* ne peut se faire de manière efficace et réaliste que dans un environnement d'aide à la décision. C'est ce qui a suscité notre intérêt pour l'opérationnalisation de l'approche par les capacités dans cette thèse.

Les outils mathématiques classiques sont insuffisants lorsqu'il s'agit de construire une évaluation des capacités des individus. Toutefois, la mathématique, comme langage de la science, a un rôle très important à jouer dans la mesure où elle regroupe un ensemble d'outils et techniques qui peuvent être utilisés pour formuler convenablement notre situation problématique. Nous considérons alors indispensable d'utiliser des modèles formels pour le mesurage de la pauvreté ainsi que l'aide à l'élaboration, l'implémentation et l'évaluation des politiques de réduction de la pauvreté. La raison est que l'utilisation de modèles formels permet d'avoir une idée claire des propriétés qu'une certaine *mesure* possède. Les modèles formels nous aident également à trouver la meilleure façon de représenter certaines observations empiriques et de comprendre si des inférences faites sur ces observations empiriques sont fondées et utiles. Enfin, ils permettent de choisir parmi différentes méthodes, procédures et approches. Le concept clé dans tout ce qui précède est la '*signifiante*'. Ainsi, pour aborder cette question d'opérationnalisation de l'approche par les capacités, nous avons introduit dans cette thèse une nouvelle méthodologie basée sur l'approche par les capacités [Sen, 1985] et l'aide à la décision [Tsoukiàs, 2008], visant la construction d'index multidimensionnel signifiante de la pauvreté (en anglais, *Meaningful Multidimensional Poverty Index*)

1.3.2 Les besoins méthodologiques

Considérons un client (ou un décideur) avec un agenda de réduction de la pauvreté incluant un certain nombre de politiques qu'il aimerait entreprendre dans une région donnée du monde. Ce client peut être, par exemple, la banque mondiale,

l'union européenne, le PNUD⁴, l'OMS⁵, le NEPAD⁶, une fondation ou un particulier, etc. qui aimerait soutenir des catégories spécifiques d'individus d'une région donnée à travers un programme d'actions spécifiques telles que l'accès pérenne aux services de base d'eau potable, d'hygiène et d'assainissement, l'accès aux services de santé et à l'éducation, la facilité d'accès au crédit, la redistribution des terres agricoles, le programme national approvisionnement en eau potable et assainissement à l'horizon 2015 (PN-AEPA horizon 2015), etc. Bien entendu, l'attribution du soutien à chaque individu devra dépendre de la classe à laquelle ce dernier appartient. Notre client est donc confronté aux problèmes majeurs suivants :

- DIFFÉRENTS TYPES DE PAUVRETÉS : C'est une erreur de parler uniquement de deux classes (la classe des "pauvres" et la classe des "non pauvres") au moins lorsqu'une perspective multidimensionnelle est considérée. Ce que nous observons en réalité sont différents types de pauvreté. Puisqu'on est différemment pauvre et de différente manière, il est crucial d'identifier des différents type de pauvreté présent dans notre base de données, puis de 'caractériser' chaque type de pauvreté afin d'y associer la politique appropriée. Ceci revêt un triple avantage : l'analyse coût-bénéfice de chaque politique, l'identification de(s) politique(s) adéquate(s) pour chaque groupe cible et la redistribution optimale des ressources disponibles. Diverses questions, quoique liées, peuvent être posées : Quel est le problème sous-jacent qui doit être traité en priorité ? Quels objectifs spécifiques doivent être poursuivis pour traiter la question des différentes pauvretés ? Qui est éligible à une certain politique ? Qui est censé bénéficier à de telles politiques ? Comment devrait-il en bénéficier ? Cette politique spécifique est-elle appropriée pour le groupe cible ? Combien devrait coûter la mise en œuvre d'une telle politique ? Que signifie lutter contre la pauvreté ?

- PLUSIEURS DIMENSIONS D'INCERTITUDE : Généralement, la table de données sur la pauvreté est très large et est formée de variables mixtes. Ainsi, l'hétérogénéité des informations doit être considérée. Celle-ci implique de considérer, outre le problème de l'imprécision horizontale (choix des dimensions), les jugements objectifs (et subjectifs) des participants au processus d'aide à la décision et l'information préférentielle du client. Il s'agit alors d'identifier les regroupements homogènes non découverts et d'établir les relations cachées entre eux afin d'extraire l'information pertinente de la base de données. On pourrait donc se poser des questions du genre : Quelles informations sont vraiment disponible et pertinent ? Which information are readily available and relevant ? Sont-elles utiles pour construire des recommandations correctes et tirer des conclusions pertinentes ? Comment accéder à l'information manquante ?

4. PNUD : Programme des Nations Unies pour le Développement.

5. OMS : Organisation Mondiale de la Santé.

6. NEPAD : New Partnership for Africa's Development.

- MESURAGE DES DIFFÉRENTS TYPES DE PAUVRETÉ : Puisqu'il existe plusieurs types de pauvreté identifiable par regroupement des individus homogènes en termes de leurs commodités, différentes perspectives doivent également être envisagées pour chaque type de pauvreté. Le revenu (ou la consommation) ne suffit pas à lui seul pour résoudre ce problème car il n'est pas toujours fiable et représentatif ; de même les seuils de pauvreté sont très discutables car ils sont généralement déterminés de façon arbitraire. Comment pouvons-nous alors mener à bien le mesurage de la pauvreté ? Étant donné que le mesurage de la pauvreté se présente alors comme un instrument de poursuite d'une certaine politique, quels éléments décrivent au mieux les conditions spécifiques de la population observée au regard de la politique qui souhaite être implémentée ? Quels éléments caractérisent au mieux chaque individu de la population étudiée comme membre d'une catégorie spécifique ? Comment pouvons-nous la mesurer ? Les personnes étant différemment pauvres, comment pouvons-nous construire des mesures reflétant parfaitement les différentes catégories de pauvreté ?

- PRÉDICTION DES CONSÉQUENCES ET VALUATION DES RÉSULTATS : Certaines politiques peuvent être s'avérer inefficace sans impact positif en moyen et long terme. Ceci peut être dû à plusieurs phénomènes telles que les incertitudes et les informations manquantes. Il est donc crucial d'élucider toutes les alternatives de manière à analyser les conséquences des diverses politiques possibles susceptibles d'être sélectionnées pour implémentation. Il s'agit aussi de rechercher et d'évaluer l'efficacité des diverses politiques possibles en mettant la meilleure évidence au centre de la recherche. On pourrait donc se poser des questions du genre : Quelles sont les alternatives potentielles ? Quelles sont leurs conséquences ? Si les résultats sont indéterminés ou incertains, comment trouver un compromis ?

Considérant la situation problématique posée par notre client, un analyste, une population donnée, nous cherchons à comprendre : Quels sont les besoins de la population ? Comment concevoir et identifier les alternatives potentielles de manière à mettre en lumière les alternatives préférables qui sont plus importantes au regard de certains objectifs et moins importantes par rapport à d'autres ? Comment décider quelle alternative politique chaque individu/ménage doit bénéficier pour améliorer ses conditions de vie standard ? Comment suivre et évaluer de telles politiques pour éventuellement les readjuster ? Dans cette thèse, nous adoptons le postulat selon lequel *la mesure de la pauvreté est le résultat d'un processus d'aide à la décision assimilable à un processus cognitif distribué*. Cette représentation de la situation problématique est un artefact qui nous permet, d'une part, de bien comprendre notre rôle dans le processus de décision et le contexte dans lequel nous construisons une mesure multidimensionnelle de pauvreté et, d'autre part, de faire comprendre au client sa position par rapport au processus de décision pour lequel il a demandé une aide. Ainsi, notre point de vue n'est pas cognitif (se rapportant à

la faculté de connaître), mais opérationnel (c'est-à-dire un instrument de poursuite d'une politique). Nous faisons l'hypothèse que les participants vont effectivement essayer de produire une représentation partagée du problème posé.

1.4 Les mesures significantes

Dans les tous premiers développement scientifique du concept de mesurage, une *mesure* est effectuée uniquement dans l'objectif d'une classification. Certains philosophes tel que [Torgerson \[1967\]](#) ne souhaitent pas appeler le résultat d'un tel processus : 'mesure'. [Roberts \[1979\]](#) considère de ce fait les avantages du "mesurage" devrait aller au-delà d'une simple classification en se fondant sur l'argument selon laquelle : *"une différence majeure entre une science bien-développée telle que les sciences physiques et d'autres sciences moins bien-développées telles que la psychologie ou la sociologie est le degré avec lequel les choses sont mesurées"*. Cet argument concernant le mesurage établit explicitement l'importance d'une 'mesure' dans le domaine des mathématiques, des sciences physiques, des sciences sociales, etc, et implicitement les décisions en découlent. La qualité d'une mesure est non seulement liée à la qualité du processus de décision ayant conduit à son développement, mais aussi des décisions qui en découlent. C'est dans cette optique que nous avons proposé dans cette thèse une nouvelle méthodologie visant la construction des mesures significantes qui aident le client et/ou le décideur dans la prise de meilleures décisions de réduction de la pauvreté à travers le ciblage des ménages les plus déprivés. Nous considérons que la mesure de pauvreté n'est pas un objet qui pré-existe et la formulation que l'on attribut à un problème ne peut pas, en général, être totalement objective et ne peut être envisagée indépendamment des rapports entre les acteurs du problème de décision et la réalité. Il est normal que cette formulation évolue au fur et à mesure de l'avancement dans le processus de décision. C'est en cela que l'aide à la décision occupe une place de choix dans la méthodologie de mesure multidimensionnelle de pauvreté que nous proposons dans cette thèse. Certes, elle ne remplace en aucun cas la phase d'apprentissage statistique visant à apprendre et à construire les politiques de décisions appropriées, mais elle l'inclut plutôt dans le processus de décision. Ce cadre établit aussi les conditions dans lesquelles nous allons opérationnaliser l'approche par les capacités de [Sen \[1985\]](#) et tirer profit de toute la richesse pour améliorer nos connaissances sur la pauvreté. Contrairement aux approches de mesure de la pauvreté existantes ([voir [Kana et al., 2011](#), pour plus de détails]), notre méthodologie de construction d'indices basée sur l'approche par les capacités et l'aide à la décision soutient le fait qu'une activité de mesure multidimensionnelle de pauvreté devrait impliquer un minimum d'insertion dans le processus de décision. Une telle activité ne doit pas se faire indépendamment des acteurs du processus de décision

et d'un cadre décisionnel bien spécifique au risque d'aboutir à une mesure de pauvreté ne reflétant pas la réalité. C'est dans cette optique que nous avons envisagé mesurer la pauvreté non seulement 'pour' le(s) décideur(s) mais essentiellement 'avec' les acteurs, les intervenants, les parties prenantes et le décideur.

1.4.1 Qu'est-ce qu'une mesure signifiance ?

Une définition plus complète du terme 'mesure' a été donnée par Mari [2003] qui atteste que "*mesure est un type spécifique d'évaluation, c'est-à-dire qu'il s'agit d'une opération visant à associer une entité d'information, le résultat de la mesure, avec l'état du système sous-jacent à la mesure en référence à une quantité donnée, le mesurande*". Nous pensons qu'une mesure de la pauvreté devrait être considérée comme un ensemble d'opérations permettant de construire un *pont* (domaine des expériences humaines subjectives) entre le *monde physique* (domaine des choses physiques) et le *monde informationnel* (domaine de la connaissance objective). Le concept de signifiance tire son origine de la théorie de mesurage [voir Suppes, 1959; Krantz *et al.*, 1971]. Roberts [1979] a présenté la 'signifiance' comme une condition essentielle pour qu'une mesure soit bien-définie dans le sens de *l'exactitude, la complétude* (ou *l'exhaustivité*) et *la rationalité*. Le point de vue de Roberts est clairement en adéquation avec la définition donnée par Stevens [1946] selon laquelle "*la mesure, dans le sens plus large, se définit comme l'affectation des nombres aux objets ou aux événements selon certaines règles*".

Toutefois, dans le cas de la pauvreté, une mesure est non seulement effectuée afin d'attribuer des nombres à des individus ou à des ménages, mais il doit aider les décideurs à prendre des décisions bien informées au sujet des politiques, des programmes et des projets. Ce point de vue se rapporte à la notion de *processus d'aide à la décision* et introduit deux conditions qui sont essentielles dans le domaine du mesurage de la pauvreté : *l'opérationnalité* et *légitimité*. Une mesure de la pauvreté est *opérationnelle* si elle peut être utilisée efficacement pour identifier les acteurs issus des univers différents et si elle peut aider les décideurs à prendre des décisions bien informées au sujet des politiques, des programmes ou des projets. De l'autre côté, une mesure de la pauvreté est *légitimée* si elle prend en compte la façon dont une recommandation finale est présentée, implémentée et perçue par les autres acteurs en dehors de son contenu précis. Notez que l'opérationnalité et la légitimité n'ont pas été explicitement définies dans ce chapitre. Les définitions données ici sont basées sur une analyse rétrospective et de reconstruction [voir Tsoukiàs, 2007] .

Donc, dans le domaine de la pauvreté, une mesure est "*signifiante*" si elle satisfait les trois conditions suivantes :

SOLIDITÉ THÉORIQUE : une mesure de pauvreté doit être *théoriquement solide*, en ce sens que les concepts utilisés pour la construire sont en adéquation avec la théorie de mesurage ;

COMPLÉTUDE OPÉRATIONNELLE : une mesure de la pauvreté doit être *opérationnellement complète*, dans le sens qu'elle est utile pour l'élaboration et l'implémentation des politiques, et aide les décideurs à prendre des décisions éclairées au sujet des politiques, des programmes ou des projets.

LÉGITIMITÉ : une mesure de la pauvreté doit être *légitime* dans le sens où, elle reflète parfaitement la perception de la société, des parties prenantes et les acteurs de la situation.

Ainsi, nous pouvons définir une mesure multidimensionnelle de pauvreté significative (MMPS) as follows :

Définition 1.4.1 *une MMPS est une mesure significative dérivée d'un processus d'aide à la décision visant à améliorer les capacités des individus et leurs conditions de vie.*

Il est important de noter que notre postulat basé sur les mesure significantes est en adéquation avec les trois positions ("Trois P") suivantes et l'approche par les capacités de Sen esquissée du point de vue de son opérationnalisation

- (P1).** Les mesures sont les propriétés intrinsèques des objets mesurés [see [Mari, 2003](#)].
- (P2).** Les mesures sont les résultats d'opérations qui préservent les relations observées entre les objets mesurés [see [Mari, 2003](#); [Roberts, 1979](#)].
- (P3).** Les mesures sont les résultats d'un processus d'aide à la décision [see [Bouysou et al., 2000](#); [Tsoukiàs, 2007](#)].

Dans une telle perspective, les mesures significantes sont un instrument crucial fournissant des solutions aux problèmes de la société tels que la santé publique et l'éducation, la lutte contre la pauvreté et l'exclusion sociale, la promotion de la redistribution efficace de la richesse, la promotion de la justice et la défense des droits de plus démunis par une gestion efficace de la richesse et des ressources.

1.4.2 Les enjeux des mesures significantes

La recherche d'une mesure efficace et plus réaliste est l'enjeu qui apparaît d'emblée lorsque l'on parle de mesure de pauvreté *théoriquement solide, opérationnellement complète et légitime*. L'un des objectifs implicites de notre nouvelle méthodologie vise à converger vers la naturalisation des résultats, c'est-à-dire aboutir à des politiques de décisions plus adéquates répondant aux préoccupations réelles des différents acteurs concernés par la problématique de la pauvreté. Cela fait allusion aux indices de pauvreté construites et aux politiques de décision adéquates qui en découlent. Notre travail de thèse vise donc à construire une mesure significative ayant quatre compétences principales : cognitive, analytique, stratégie politique, suivi et évaluation.

1.4.2.1 La compétence cognitive

La littérature sur la pauvreté, aussi bien dans les pays en développement que dans les pays développés, suppose plus souvent qu’il existe intrinsèquement deux classes sociales : *les pauvres* et *les non-pauvres*. Cette supposition est une erreur monumentale d’autant plus que l’appellation *pauvre* ou *non-pauvre* n’a objectivement aucun sens. La “vérité” est que le pauvre est celui à qui il manque quelque chose et cette chose manquante varie au niveau d’un indicateur à un autre dans une région donnée. De ce fait, toute définition de la pauvreté devient mensongère lorsqu’elle porte une connotation objective. Les méthodes d’analyse des données sont donc sollicités pour identifier *automatiquement* les différents types de pauvreté présents dans nos données. Notons que, l’idée de base de l’analyse des données est de réduire au maximum la subjectivité et les a priori : “*le modèle doit suivre les données et non l’inverse*”. La compétence cognitive revêt donc une importance cruciale en ce qu’elle permet d’appréhender la situation de pauvreté et d’identifier les différents groupes d’individus différemment pauvres. Elle précède une analyse des données dont le but essentiel est la mise en relief des relations existantes entre les individus, entre les attributs qui les caractérisent, et entre les individus et les attributs. Donc, la compétence cognitive permet d’effectuer une classification non supervisée (ou *clustering* en anglais) des individus pour identifier les types de pauvreté.

1.4.2.2 La compétence analytique

L’étude de la pauvreté conduit à tirer parti des données sur la pauvreté d’un pays (d’une région ou d’une communauté) pour la synthétiser, pour servir de base à un processus de décision, de reconnaissance dans le but d’appréhender la nature des phénomènes sous-jacents aux données. L’étude cognitive de la pauvreté permet d’avoir une situation globale des individus les uns par rapport aux autres et de former les classes de pauvreté. La compétence analytique conduit à la description et la caractérisation des différentes classes explicitées pendant la phase cognitive. La phase analytique consiste donc en l’identification et la définition des facteurs déterminant chaque classe de pauvreté. Celle-ci passe par la mise en relief des relations existantes entre les classes, entre les classes et les attributs, puis la mesure de l’importance relative de chaque classe dans cette répartition. Elle est cruciale dans la mesure où elle permet de *comprendre* chaque classe à partir des caractéristiques des individus, d’identifier les éventuelles déprivations et de déterminer les classes cibles par rapport une politique que le décideur aimerait entreprendre.

1.4.2.3 La compétence stratégie politique

Associer chaque ménage à une classe permet d'une part de mieux décrire le cadre précis de la cible d'une certaine politique et de mesurer son efficacité dans les temps ; et d'autre part, d'avancer des hypothèses alternatives d'action si le modèle se révèle très éloigné par rapport à la réalité observée ou les priorités du décideur en terme de politiques. En d'autres termes, notre approche veut proposer une méthodologie de mesurage qui est opérationnelle pour une certaine politique et pour un certain cadre décisionnel observé tout en respectant le cadre formel de la théorie du mesurage (éclairage de la décision). Il s'agit aussi dans cette compétence de dire comment les indices construites seront traduites en politiques concrètes en terme d'action, de programmes et de projets.

1.4.2.4 La compétence suivi et évaluation

Pour juger l'efficacité d'une stratégie de lutte contre la pauvreté (S.L.P.) en matière de réduction de la pauvreté, l'analyste doit élaborer un système de suivi et d'évaluation permettant de suivre les indicateurs clés dans le temps et l'espace pour voir s'ils changent du fait de la stratégie. Ensuite, l'analyste doit pratiquer des estimations rigoureuses et sélectives dans le but d'une évaluation de l'impact des interventions ou des politiques de décision qui sont des composantes clés de la stratégie. L'approche développée dans cette thèse est conçue de manière à faciliter une évaluation rigoureuse des impacts et la mise en oeuvre d'un système de suivi de la pauvreté. Notre approche opérationnelle permet également de déterminer l'adéquation d'une S.L.P. en matière de réduction de la pauvreté et à les renforcer au besoin.

1.5 Objectif générale de la thèse

L'objectif générale de cette thèse est de construire un modèle d'évaluation globale de la pauvreté destiné à la conception d'indices de pauvreté innovants et intégrables dans des systèmes d'aide à la décision. Ces indices de pauvreté découlent des mesures significatives construites dans le but de mettre en évidence les différentes formes de pauvreté observées et le meilleur ciblage des populations pauvres en vue de la rationalisation des politiques de réduction de la pauvreté. Le modèle devra être suffisamment générique et stable pour pouvoir être, par la suite, adapté et transposé à plusieurs communautés, régions, pays. Ce modèle vise entre autre à établir un langage et une méthodologie communs à l'ensemble des individus (ou ménages) d'une société donnée.

La construction d'une mesure significative multidimensionnelle de la pauvreté nécessite à la fois une étude méthodologique (comment les indicateurs peuvent

être signifiants) et une étude sur les politiques pour lesquelles cette mesure peut être utile. De ce point de vue, le modèle permettra alors de répondre à la problématique de la pauvreté, aux questions d'élaboration et d'implémentation des politiques. Il devra être à la fois *descriptif*, de par sa pertinence lorsqu'il s'agira de retrouver une "*description*" du problème qui nous permettrait de découvrir une *meilleure décision* dans le cadre des politiques proposées par notre client ou décideur (notamment le NEPAD, l'Union Européenne et la Banque Mondiale) et *prescriptif*, de par son utilisation pour servir de base à un outil de diagnostic-conseil auprès des organismes internationaux, organismes régionales et sous-régionales, en adoptant une approche *hybride* et *constructive* basée sur la segmentation et les méthodologies d'aide multicritère à la décision. Cet objectif global se décline autour de cinq préoccupations principales :

- a) LA SIGNIFIANCE : en ce sens que les concepts utilisés pour construire la mesure de la pauvreté sont en adéquation avec la théorie de mesurage.
- b) L'OPÉRATIONNALITÉ : en ce sens que la mesure de la est utile pour l'élaboration des politiques, l'implémentation des politiques et elle aide les décideurs à prendre des décisions bien-informées au sujet des actions concrètes, les programmes ou les projets.
- c) LA LÉGITIMITÉ : la mesure de la pauvreté doit être légitimée dans le sens où, il devrait refléter la perception de la société, des acteurs et des parties prenantes.
- d) LA COMPLEXITÉ : elle implique la prise en compte de la diversité humaine, la prise en compte de la diversité des agents vecteurs de la pauvreté, la prise en compte des différentes facettes de la pauvreté et de son caractère flou.
- e) LA ROBUSTESSE : les conclusions et les recommandations finales doivent être robustes, c'est-à-dire être supportées par les données fournies et validées par des scénarios après plusieurs simulations.

Première partie

État de l'art

Analyse et mesure de la pauvreté

2.1 Introduction

Le mesurage de pauvreté peut être perçu comme un outil fondamental dans la lutte contre la pauvreté, considérée comme un fléau de l'humanité. Depuis de nombreuses années, la recherche d'une mesure de pauvreté faisant l'unanimité obsède les États, les décideurs et la société. Avec l'émergence des problèmes liés à la pauvreté et l'étrange persistance de ce fléau, les observatoires, les études et les recherches sur les causes de la pauvreté se sont multipliés, mobilisant beaucoup de philosophes, de sociologues, d'économistes et les mathématiciens. Cette mobilisation des différents chercheurs a contribué à l'élaboration de nombreux concepts et théories.

Le but de ce chapitre est de présenter une revue de littérature sur la genèse du concept de '*pauvreté*' et de '*mesure de la pauvreté*', et des différentes approches établies en vertu d'une quête de soutien scientifique dans la lutte contre la pauvreté. La perspective générale visée au travers de cette revue de la littérature est de montrer comment le concept de la pauvreté a évolué dans le temps, les difficultés quant à sa conception et sa mesure. L'objectif final étant bien-sûr de répondre à la question : comment évoluer dans la conceptualisation et la mesure de pauvreté ? En effet, si la pauvreté doit être mesurée, c'est parce que de telles mesures devront être utilisées pour élaborer des décisions appropriées afin de construire des politiques adéquates de réduction de la pauvreté. Sous une telle perspective, nous allons introduire le concept de '*signifiance*' d'une mesure de la pauvreté.

Une méthode commune utilisée pour mesurer la pauvreté est basée sur le revenu ou la consommation lequel indique qu'un individu ou un ménage est considéré comme '*pauvre*' si son revenu ou son niveau de consommation est inférieur à un certain niveau minimum (appelé *ligne de pauvreté* ou *seuil de pauvreté*) nécessaire pour satisfaire ses besoins fondamentaux. Cependant, la diversité des causes de la pauvreté et la complexité de ses conséquences oppose une résistance logique à une telle approche. En effet, la notion de pauvreté peut être expliquée en sociologie, mais les outils traditionnels mathématiques ne peuvent pas appréhender le niveau réel de ce fléau. Car son aspect multidimensionnel, son caractère flou et les substitutions possibles entre les différentes formes qu'elle peut prendre, font en sorte que les recherches visant à construire une mesure adéquate de pauvreté et

la surveillance de ses facteurs explicatifs soient plus complexes. La pauvreté étant multidimensionnelle, son mesurage doit intégrer l'aspect multidimensionnel et en même temps permettre d'explorer les déprivations individuelles et d'identifier les différentes formes de pauvreté existante.

Ce chapitre est organisé comme suite : la façon dont la pauvreté a été perçue dans l'histoire et les fondations théoriques de la mesure de la pauvreté sont explorées dans la Section 2.2. Dans la Section 2.3 présentera nos notations, quelques définitions formelles de la mesure de la pauvreté et la variété des approches de mesure multidimensionnelle de la pauvreté existantes dans la littérature.

2.2 Un peu d'histoire

2.2.1 La pauvreté : origines et fondements conceptuels

La pauvreté n'est pas un phénomène datant de nos jours. Elle a existé pendant des siècles et continue de faire rage aujourd'hui dans de nombreux pays du monde. Le concept de pauvreté prend ses racines dans l'éthique sociale, qui peut être considérée comme un élément central de la philosophie politique. Les facettes de la pauvreté tout au long de l'histoire sont bien connues et visibles, de la société athénienne à partir de laquelle Socrate, Platon, Aristote et Xénophon [Sanderson, 2004] ont argué au sujet de la pauvreté et la richesse, à notre époque où les parlementaires et les organisations internationales traitent des portefeuilles d'actions contre la pauvreté. Depuis la fin du moyen âge jusqu'à l'aube de la révolution industrielle, la lutte contre la pauvreté s'est basée sur différentes approches conceptuelles sans l'apparition d'une 'vraie' solution à ce fléau.

De 1500 à 1700, la pauvreté n'est pas dépouillée de sa dimension religieuse. Cependant, cette vision théologique du pauvre est influencée par des périodes de changement. En Europe, elle est perçue comme un *désordre civil*, créé par la mendicité ; *désordre moral*, causé par la paresse et l'oisiveté des mendiants. En Angleterre, le Parlement anglais et la reine Elizabeth I (1533-1603) ont travaillé ensemble pour adopter des lois [cf. Marshall, 1985; Slack, 1990] conçues pour lutter contre la pauvreté afin de réduire le désordre civil et moral. Dès cet instant, les pauvres (aussi appelés *malheureux*) sont divisés en deux groupes : les "méritants" : les malades, les handicapés, les veuves, les orphelins et les anciens économes ; et les "non méritants" : les délinquants, les mères célibataires, les vagabonds, les chômeurs et les vieux sans épargne. Ainsi, à partir de 1597, une série de lois, collectivement appelée "*Poor Law*", concernant la pauvreté en Angleterre sont passées afin de fournir une assistance du gouvernement à ceux qui ne pouvaient pas s'en sortir eux-mêmes, c'est-à-dire les méritants, y compris les personnes âgées et les jeunes enfants. Cette loi devint, pendant plus de 200 ans, la première législation importante permettant

de traiter la question des pauvres et des défavorisés. Elle devint également la base d'autres législations traitant la question d'assistance aux pauvres au niveau colonial et permettant au gouvernement d'établir des programmes d'apprentissage destinés aux enfants pauvres, de développer des *maisons de travail* ('workhouses') pour ceux qui étaient en mesure de travailler ("les bons pauvres") et des *maisons de correction* pour les vagabonds ("les mauvais pauvres"). En 1662, la recrudescence des "mauvais pauvres" parmi ceux considérés comme "bons pauvres" conduisit à l'avènement d'une nouvelle loi en Angleterre appelée "*Law of Settlement and Removal*". Les autorités visaient à travers cette loi à aider uniquement les citoyens pauvres locaux ("bons pauvres") et à expulser de la juridiction toute personne qui pourrait devenir dépendante de l'aide ("le mauvais pauvre"). Cette loi amena les autorités à codifier les causes de la pauvreté en évaluant les gens par la *vraisemblance d'être pauvre* et en sanctionnant sévèrement les "mauvais pauvres". Cela représente jusqu'à présent l'une des premières lois au monde portant sur la détermination des conditions d'admissibilité à recevoir de l'aide ("*residency requirements*"). Cette loi est restée en place longtemps après le règne d'Elizabeth, fournissant ainsi une base pour l'élaboration des politiques du gouvernement destinées à l'amélioration des conditions de vie des pauvres jusqu'au 19^e siècle [Prothero, 1894]. Telle une question de faits, de telles politiques ont été conçues plus en terme de la réduction de l'impact de visibilité des pauvres dans les sociétés riches de cette époque, plutôt que dans le but pour s'attaquer aux problèmes réels des pauvres et aux raisons pour lesquelles les gens devenaient pauvres.

Au début du 17^{ème} siècle, alors que certains gouvernements mettent en place des *centres d'assistance aux pauvres* dans leurs territoires, d'autres, telle que la Nouvelle-Écosse, choisissent d'adopter pleinement les lois anglaises "Poor Laws" dans leur pays. En 1795, le système de Speenhamland [cf. Marshall, 1985; Slack, 1990] établit le premier "seuil de pauvreté" en se basant sur le prix du pain et le nombre de personnes prises en charge dans une famille de travailleurs ; définissant ainsi les subventions fournies lorsque les salaires étaient en dessous du seuil de pauvreté. Le système de Speenhamland [voir Marshall, 1985; Slack, 1990] était perçu comme une forme d'assistance accessible à tous, destinée à atténuer la pauvreté rurale, à la fin du 18^{ème} siècle et au cours de début du 19^{ème} siècle. Toutefois, les 17^{ème} et 18^{ème} siècles sont marqués par les premières réflexions philosophiques sur le concept de *pauvreté* et de le concept de *mesure de la pauvreté* à la source de l'économie du bien-être. Bentham [1789], précurseur des réflexions philosophiques sur le concept de pauvreté, part du principe que les individus ne conçoivent leurs intérêts que sous le rapport du *plaisir* et la *peine* pour proposer une nouvelle doctrine qui portera le nom d'*utilitarisme* dès 1781. L'idée consiste à déterminer scientifiquement la valeur qui maximise le bonheur individuel (exprimée par le surplus de plaisir sur la peine) en utilisant des règles précises telles que la quantité de plaisir et de peine générée par nos diverses actions. Les résultats préconisent le

retour au travail comme moyen principal de la lutte contre la pauvreté et dans le même temps contre un facteur d'entropie sociale. On demande alors à l'État de contractualiser les rapports humains, sans contrainte extérieure, afin de garantir à la société un '*plus grand bien pour le plus grand nombre*'. C'est dans ce contexte que le concept d'*utilité cardinale* [voir Ellsberg, 1954; Stigler, 1950] a été introduit et défini en termes de *bonheur* ou de *plaisir* [Bentham, 1789], et de satisfaction des désirs [Sidgwick, 1874]. Dans ce contexte, il s'agit alors *de maximiser le plaisir et la satisfaction d'un grand nombre (bien-être global), tout en minimisant la peine et la souffrance de personnes* vivant dans une communauté donnée. D'autre part, le prédicateur et mathématicien écossais Thomas Chalmers [cf. Young et Ashton, 1956] recommanda en 1819 *d'aider les pauvres à s'aider eux-mêmes*. Des fondations philanthropiques privées sont à cet effet développées pour aider à répondre aux besoins économiques des populations pauvres. Thomas Chalmers prescrit donc d'améliorer le sort de ses semblables par de multiples moyens, et ceci de manière désintéressée. Il organise un système de bénévoles qui se rencontrent individuellement et régulièrement avec les personnes défavorisées afin de les encourager et de les former. C'est ainsi que les premiers programmes sociaux basés sur un système de bénévoles pour fournir une aide économique d'urgence et une assistance "spirituelle" aux pauvres ont émergé.

Le 19^{ème} siècle, marqué par l'industrialisation, sera caractérisé par une image négative de la pauvreté. En 1834, une nouvelle loi, appelée "*Poor Law Amendment Act (PLAA)*" fût promulguée en Angleterre en remplacement de la "*Poor Law*" de 1601. Cette nouvelle loi mit l'accent sur l'*autonomie*, et sa principale caractéristique fût la création d'hospices de dissuasion. L'aide publique n'est plus considérée comme un droit, et le gouvernement n'est plus tenu à prendre en charge les chômeurs. Le principe de "*less eligibility*" est appliqué. Ce principe stipule que : "*un bénéficiaire de l'aide ne peut jamais recevoir autant qu'on ne paye un salarié ou celui qui a travaillé*". Toute personne sollicitant une aide sociale devrait d'abord avoir intégré la "*workhouse*". Si un homme a demandé une aide sociale, alors toute sa famille devrait également avoir intégré la "*workhouse*", et l'aide à lui attribuée sera définie en conformité avec les *règles de séparation*. En 1847, la "*Poor Law Commission*", qui était chargée de l'attribution des aides, a abolié la *Poor Law* en raison des abus qui ont eu lieu à cette époque, notamment à cause du scandale "*Andover Workhouse Scandal*"¹ exposé par les journalistes et les politiciens en 1845. Toutefois, la pauvreté est encore perçue comme un phénomène de masse que Marx [1850] décrit sous les traits du prolétariat (la *classe sociale des prolétaires* également connue sous le nom des *travailleurs pauvres*). Le mot *prolétaire* désigne à l'origine un citoyen romain qui n'a que ses enfants comme richesse. Les prolétaires sont des travailleurs manuels de la grande industrie et forment la classe la moins

1. See http://en.wikipedia.org/wiki/Andover_workhouse_scandal, August 2011.

considérée du *civitas* (ensemble des citoyens), constituée de ceux qui ne peuvent s'acheter aucune pièce d'armure et qui ne possèdent le droit de vote qu'en théorie. C'est la dernière classe sociale. Bien sûr, l'intérêt du prolétaire est d'obtenir le plus possible de son travail, tandis que réciproquement le propriétaire des moyens de production cherche à minimiser ce coût. Le prolétariat a donc des intérêts exactement contraires à ceux du bourgeois (ou capitaliste). D'où un conflit connu sous le nom de "*la lutte des classes*" [Marx, 1850, 1887; Milano, 1992] survient entre le prolétariat et la bourgeoisie. Le pauvre redevient un "*mauvais pauvre*", inutile parce que interchangeable, utile seulement lorsqu'il est vertueux. Le pauvre appartient alors à une classe sociale (le prolétariat) se définissant par l'obligation de se nourrir (ou de vivre) en vendant sa force de travail à la classe antagoniste (bourgeois ou capitaliste) qui dispose du capital et des moyens matériels de production. Les différentes phases de la révolution de 1848 en France [cf. Marx, 1850] reflètent le caractère nouveau de la lutte des classes telle qu'elle apparaît dans le milieu du XIXe siècle. En 1889, Bismarck, chancelier d'une nouvelle Allemagne unifiée, présente le premier système national d'assurance maladie. Les questions d'égalité entre les individus deviennent une préoccupation sociale et occupent une place importante dans le discours politique. Mill [1848] refuse cette vision de Bentham de la société et pose le problème de la distribution des droits de propriété. Mill [1871] part du constat qu'il existe des inégalités de conditions dans la distribution des droits de propriétés (une partie des richesses revient à une minorité) et propose de poursuivre uniquement la minimisation de la souffrance (physique et morale) afin de tenir compte du bien-être des individus seulement lorsque celui-ci est négatif, c'est-à-dire qu'il est sans un *mal-être*. L'idée de Mill a été enrichie pour aboutir au recours à l'État pour réajuster la distribution des richesses. Toutefois, la *doctrine utilitariste* [voir Bentham, 1829, 1831; Mill, 1871; Sidgwick, 1874] est restée dominante jusqu'à la fin du 19ème siècle.

Le 20ème siècle a été marqué par la poursuite du processus d'industrialisation, le progrès technologique et l'augmentation générale du niveau de vie qui ont progressivement amélioré la situation des "travailleurs pauvres" jusqu'à la période d'après-guerre connu sous le nom de "Thirty Glorious Years" (1945-1975). Ce siècle est principalement marqué par un mouvement de contestation contre la *doctrine utilitariste* dans différents pays. Les critiques découlent des assertions de Bentham et Mill arguant que l'utilité est une notion psychologique exprimée par une seule et même personne. Il devient alors difficile de classer les états psychologiques qui révèlent des conceptions de satisfaction personnelle et d'exprimer l'utilité sur une échelle de valeurs communes. Pour résoudre ce problème, Pigou [1920] propose de monétiser la fonction d'utilité. Il mesure alors la satisfaction en se basant sur le montant d'argent qu'une personne est prête à offrir pour obtenir un certain panier de biens. De ce point de vue, il devint possible de comparer les niveaux d'utilités entre individus. Les résultats obtenus par Pigou [1920] sont à l'origine

de “l'économie du bien-être” (“*Welfare Economics*”) et de la notion d'*utilité ordinaire* basée sur le classement des préférences des individus [Robbins, 1938; Pareto, 1927]. Ils préconisent notamment de maximiser le bien-être social (bien-être collectif) en maximisant une fonction d'agrégation des utilités individuelles, c'est-à-dire faire en sorte que les gens soient prêts à payer plus pour consommer. Les travaux de Pareto [1896-97] soulèvent la question de l'allocation optimale des ressources qui se présentent sous formes de deux questions clés : Comment améliorer le sort d'une personne sans nuire à d'autres personnes ? Est-il possible de augmenter la satisfaction de tous les individus en même temps ? Il s'agit de trouver l'optimum de Pareto (*Equilibrium State*), à savoir, une allocation efficace des ressources de manière à maximiser l'utilité sociale quelque soit la répartition des richesses entre les individus. Des réponses partielles à ces deux questions seront fournies par la “théorie du choix social” et mettra en évidence certaines théories économiques actuelles qui étudient en particulier les questions de pauvreté. Le problème consiste alors à proposer une relation de préférence collective (économique du bien-être social) à partir des préférences individuelles, c'est-à-dire, de construire une règle d'agrégation des fonctions d'utilité individuelles en une fonction de bien-être social [see Arrow, 1951; Bergson, 1938; Sen, 1970]. Avec la publication de Rawls [1971], les choix politiques en matière de justice sociale se façonnent peu à peu dans la société. Rawls [1971] remet en cause l'*utilitarisme* dans son ouvrage phare [Rawls, 1971] et propose un postulat “*égalitariste*” qui évoque qu'une société est juste lorsqu'elle distribue, de manière équitable, un ensemble de ressources, permettant à l'individu responsable, de mener l'ensemble des plans rationnels qu'il s'est fixé. Une autre école propose de dépasser à la fois le *welfarism*² et l'*égalitarisme* pour se tourner vers l'égalité des opportunités, dont l'objectif est d'offrir aux citoyens les mêmes opportunités d'améliorer leur qualité de vie [Sen, 1979a].

2.2.2 Les fondements du mesurage de la pauvreté

Les premiers travaux³, considérés comme scientifiques, traitant de la question de la pauvreté et de ses manifestations économiques et sociales remontent aux études de Charles Booth [1892, 1894] et Seebohm Rowntree [1901] sur la situation des ménages les plus pauvres à Londres et au Royaume-Uni. La conception selon laquelle le revenu est la seule variable fiable pour mesurer la pauvreté était fortement ancrée parmi les économistes. Tandis que Booth [1892] développait une méthodologie purement exploratoire sans propositions plus concrètes pour éradiquer la pauvreté (autre que l'allocation de pensions de solidarité envers les plus

2. Le *welfarism* fait référence au bien-être économique, où la théorie du bien-être repose sur l'idée selon laquelle la qualité d'une situation peut être jugée entièrement en fonction de la qualité des utilités prévalant dans cette situation [see Sen, 1979c,a].

3. Pour plus de détails et autres références sur le sujet, consulter Fusco [2005]; Bertin [2007].

démunis), Rowntree [1941] utilisait une méthode plus scientifique basée sur une évaluation des ressources à disposition des ménages pour aboutir à des résultats statistiques lui permettant de proposer des pistes de réformes économiques en faveur d'un soutien financier aux plus pauvres, grâce à la mise en place d'un salaire minimum et de prestations sociales pour les familles nombreuses.

Les années 1950 sont marquées par les *théories de la croissance* stipulant que toutes les économies devront tendre à terme vers un niveau de développement tel que la croissance économique profitera à tous [Aghion et Bolton, 1997]. Cette année donna naissance à la notion de *pauvreté macroéconomique*, qui est une situation au cours de laquelle la croissance du Produit Intérieur Brut (P.I.B.) n'est pas suffisante pour que toutes les classes d'une population profitent de ses fruits. De ce fait, lutter contre la pauvreté revenait simplement à construire des mécanismes de "maximisation" de la croissance économique.

Les années 1960 marquent une progression dans la compréhension de la pauvreté et le *paradigme monétaire* commence à être considérablement critiqué. Bien que la pensée économique reste influencée par le modèle libéral qui prône une allocation efficace des ressources lorsque les mécanismes de marché sont facilités, les sociologues [Abel-Smith et Townsend, 1965; Runciman, 1966] remettent en cause le paradigme monétaire en partant d'une vision relative de la pauvreté. Les publications de la fin des années 1960 témoignent de l'évolution de la pensée économique basée sur le paradigme monétaire vers des considérations plus sociales de la pauvreté [Bauer, 1967; Van Praag, 1968; Seers, 1969] et l'introduction du *paradigme multidimensionnelle* [Abel-Smith et Townsend, 1965; Runciman, 1966].

Entre la fin des années 1960 et le début des années 1970, la croissance économique joue encore un rôle central. Mais, seule une partie de la population, la classe des personnes aisées, profite des fruits de la croissance au détriment du reste de la population. Il convient donc d'agir directement sur les plus pauvres par une redistribution des fruits de la croissance économique. Toutefois, les années 1970 sont marquées par l'avènement de la crise de la dette dans les pays en voie de développement, les chocs pétroliers et l'abandon du système des taux de change fixes des Institutions de Bretton Woods (I.B.W.). Cette conjonction d'événements négatifs et déstabilisateurs a entraîné une envolée de l'inflation et un retour du chômage qui a dès lors commencé à accélérer inexorablement. Les situations d'exclusion et de précarité se sont accrues dans certaines classes sociales, notamment la classe sociale des défavorisés. Il en découle que le développement d'un système élargi de redistribution et d'assurance sociale visant à réduire les inégalités et la pauvreté monétaire n'est pas suffisant pour éliminer les frustrations économiques et sociales. La recherche d'une stabilisation macroéconomique conduit à la mise en place de Programmes d'Ajustement Structurel (P.A.S.) dont l'objectif est de rétablir les conditions de remboursement des dettes externes contractées par les pays en développement. Il s'agit d'un ensemble de dispositions dont certaines agissent sur

la conjoncture, d'autres sur les structures et qui résultent d'une négociation entre un pays endetté et le F.M.I. (ou la banque mondiale) pour améliorer le fonctionnement économique du pays (le F.M.I. conditionnant son aide à la mise en place de réformes pérennes). De ce fait, un ensemble de mesures d'ajustement structurel sont élaborées obligeant les pays endettés à mettre en œuvre un processus de libéralisation de l'économie, de privatisation des grandes entreprises nationales et de réformes du système monétaire, bancaire et financier. Ces P.A.S. sont très rapidement voués à l'échec à cause de l'absence de considérations des plus pauvres lors de leurs mises en œuvre, et de l'accélération du processus de mondialisation qui a ouvert les portes des marchés des pays du Sud aux investisseurs occidentaux [cf. Bertin, 2007].

Les années 1980 sont marquées par l'échec des P.A.S. [cf. Zattler, 1989; Shah, 2010] qui a contribué à augmenter la pauvreté de certains pays sujet à ces programmes. La pauvreté, surtout dans les pays africains, prend des proportions sans précédent, tant en extension qu'en intensité. Le ton est donné dans une série de critiques [cf. Cornia *et al.*, 1987; Commonwealth, 1989; Watkins, 1995] remettant en cause l'approche exclusivement économique des programmes d'ajustement structurel pour réclamer un ajustement à visage humain. À la suite des critiques encourues, le P.N.U.D, la Banque Africaine de Développement (B.A.D.) et la Banque mondiale [World Bank, 1990] lancent dès 1989 le programme "*dimension sociale de l'ajustement*" pour intégrer dans ses programmes un volet visant l'atténuation des conséquences des ajustements sur les catégories de population les plus pauvres. Cette volonté de considérer l'ensemble des effets socio-économiques négatifs des programmes macro-économiques d'ajustement sur les populations des pays concernés reste cependant marquée par le libéralisme économique, auquel on adjoint des filets de protection sociaux prenant en compte les effets non désirés sur les populations [Morduch et Sharma, 2002].

En 1990, le rapport du P.N.U.D. [PNUD, 1990–2006] introduit l'approche par le *développement humain* et pose explicitement le problème de la pauvreté. Cette dernière n'est plus seulement un état de privation monétaire, mais fait référence à l'ensemble des possibilités et des choix offerts aux individus. Le rapport de 1994 revient sur la question en reliant la pauvreté à la croissance économique avec des considérations sociales, politiques ou éthiques. En 1996, le Comité d'aide au développement de l'O.C.D.E. s'engage à éradiquer l'extrême pauvreté dans le monde à l'horizon de l'an 2015 tandis que l'O.N.U. proclame "*l'année internationale pour l'éradication de la pauvreté*" et déclare en même temps la période 1997-2006 comme "*décennie internationale pour l'éradication de la pauvreté*". L'année 1996 marque aussi une nouvelle entrée en scène du F.M.I, conjointement avec la Banque mondiale dans l'initiative P.P.T.E. (Pays Pauvres Très Endettés). L'initiative P.P.T.E. consiste à accorder une assistance exceptionnelle aux pays éligibles qui mettent en œuvre des politiques économiques saines, pour les aider à rame-

ner leur niveau d'endettement à un niveau tolérable, de telle sorte qu'ils puissent par la suite assurer le service de la dette par leurs recettes d'exportation et l'aide extérieure. En Septembre 1999, le Comité intérimaire du F.M.I. souscrit au remplacement de la Facilité d'ajustement structurel renforcée (FASR) par une facilité nouvelle, la Facilité pour la réduction de la pauvreté et la croissance (FRPC), qui vise à faire des efforts de réduction de la pauvreté des pays membres à faible revenu un élément décisif et plus explicite d'une stratégie économique renouvelée orientée sur la croissance. La nouvelle facilité a pour objectif de *“soutenir les programmes destinés à renforcer de manière substantielle et continue la position de la balance des paiements [des pays membres à faible revenu éligibles] et de contribuer à une croissance durable, qui conduise à des niveaux de vie plus élevés et à une réduction de la pauvreté”*. C'est ainsi que le F.M.I. établit explicitement le lien entre l'ajustement et l'endettement. Toutefois, le clivage entre les tenants d'une pauvreté monétaire (fondée sur un manque de revenu permettant de vivre décemment) et les tenants d'une pauvreté multidimensionnelle (qui s'exprime à travers un ensemble d'indicateurs socio-économiques) demeure tout au long des années 1990, et se poursuit jusqu'au début des années 2000.

Le début des années 2000 est marqué par un triple mouvement vers une universalité du discours sur la pauvreté avec la promotion des droits universels [[World Bank, 2001](#)]; une approche plus locale de la pauvreté avec les approches qualitatives et participatives; et la Déclaration, approuvée par 189 pays, engageant les nations signataires à participer à un nouveau partenariat mondial visant à réduire l'extrême pauvreté et fixe une série d'objectifs devant être atteints en 2015 et appelés *Objectifs du Millénaire pour le Développement* (O.M.D.). Ce mouvement ouvre un consensus entre le P.N.U.D. et la Banque mondiale. Ces deux organisations considèrent que la pauvreté résulte d'un manque d'accès aux actifs, d'une croissance économique insuffisante ou inappropriée, et d'une mauvaise gouvernance. Cette dernière notion a pris une importance croissante dans le discours du P.N.U.D. et de la Banque Mondiale sur la pauvreté et le développement, où l'accent a été mis sur les structures politiques et les institutions des pays, ainsi que sur la nature et sur le comportement de l'État. L'accent placé par les organisations internationales sur la dimension politique ou institutionnelle de la pauvreté se retrouve, d'une part, dans l'importance accordée dans leurs analyses à la liberté de choix des individus, et, d'autre part, dans les politiques qu'elles proposent pour lutter contre la pauvreté [[Bénicourt, 2005](#)]. La Banque mondiale s'approprie le discours sur la pauvreté développé conjointement par Sen (et par le P.N.U.D.) et lui accorde une place privilégiée en incluant une citation de Sen dans son rapport en 2001 : *« Les pauvres vivent sans les libertés d'action et de choix fondamentales que les personnes mieux loties considèrent comme des choses dues [Sen, 1999]. Elles manquent souvent de nourriture et de logements adéquats, d'éducation et de santé. Ces privations les empêchent de mener le type de vie que tout le monde valorise »* [[World Bank, 2001](#)].

La Banque Mondiale reprend donc l'idée selon laquelle la pauvreté limite la « *liberté de mener une vie de valeur* ». Elle rejoint ainsi la perspective en termes de développement humain du P.N.U.D. sans toutefois proposer un concept du même type dans les rapports sur le développement dans le monde.

2.3 Mesurage multidimensionnel de la pauvreté

2.3.1 Notations and définitions

Nous adopterons les notations suivantes tout au long de cette première partie. Soit $\Omega = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, l'ensemble des individus⁴ d'un pays (région, communauté, ville, etc.) donnée avec $|\Omega| = \text{card}(\Omega) = n$. Soit $X = X_1 \times X_2 \times \dots \times X_m$, le produit cartésien des $X_j \subseteq \mathbb{R}^+$ avec $j \in \{1, 2, \dots, m\}$ et $m \geq 2$. $X \subseteq \mathbb{R}_+^m$, où \mathbb{R}_+^m est l'orthant non-négatif de m -espace euclidien \mathbb{R}^m . Traditionnellement, une distribution multidimensionnelle d'une population de n individus est donnée par une matrice des réponses $n \times m$ suivante :

$$X = \begin{matrix} & \mathbf{x}^1 & \dots & \mathbf{x}^j & \dots & \mathbf{x}^m \\ \mathbf{x}_1 & \left[\begin{array}{cccc} x_1^1 & \dots & x_1^j & \dots & x_1^m \end{array} \right. \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{x}_i & \left. \begin{array}{cccc} x_i^1 & \dots & x_i^j & \dots & x_i^m \end{array} \right. \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{x}_n & \left. \begin{array}{cccc} x_n^1 & \dots & x_n^j & \dots & x_n^m \end{array} \right. \end{matrix} \quad (2.1)$$

où x_i^j est l'évaluation du i ème individu sur le j ème attribut. Soit z^j le seuil de pauvreté ou minimum acceptable pour l'attribut $j \in \{1, 2, \dots, m\}$ et $\mathbf{z} = (z^1, z^2, \dots, z^m)$ le seuil de pauvreté multidimensionnel avec $\mathbf{z} \in Z$ et $Z \subseteq \mathbb{R}_+^m \setminus \{0\}$. Soit $\mathbf{x}_i = (x_i^1, x_i^2, \dots, x_i^m)$, l'évaluation multidimensionnelle de l'individu $i \in \{1, 2, \dots, n\}$ et $\mathbf{x}^j = (x_1^j, x_2^j, \dots, x_n^j)^T$ (où T représente l'opérateur *transposé*), les évaluations des n individus sur l'attribut $j \in \{1, 2, \dots, m\}$. Soit D_j , l'ensemble des individus déprivés sur l'attribut $j \in \{1, 2, \dots, m\}$. Soit $f(z^j, x_i^j)$ la déprivation de l'individu i sur l'attribut j dans la situation X . Le profil de déprivation sur l'attribut $j \in \{1, 2, \dots, m\}$ est donné par :

$$\mathbf{f}^j(\mathbf{z}, X) = \mathcal{G}(f(z^j, x_1^j), \dots, f(z^j, x_n^j)) \quad (2.2)$$

$\mathbf{f}^j(\mathbf{z}, X)$ est une application de $Z \times X \rightarrow \mathbb{R}_+$ qui définit la procédure d'agrégation des évaluations individuelles sur un attribut $j \in \{1, 2, \dots, m\}$ en un *indicateur cardinal* de pauvreté notée π^j .

4. En fonction des cas étudiés, l'unité statistique pourra être un ménage, un enfant, un agriculteur, etc.

Le profil de déprivation de l'individu $i \in \{1, 2, \dots, n\}$ est donné par :

$$\mathbf{f}_i(\mathbf{z}, X) = \mathcal{F}(f(z^1, x_i^1), \dots, f(z^m, x_i^m)) \quad (2.3)$$

$\mathbf{f}_i(\mathbf{z}, X)$ est une application de $Z \times X \rightarrow \mathbb{R}_+$ qui définit la procédure d'agrégation des multiples dimensions en un *indicateur composite* de la pauvreté noté π_i .

\mathcal{G} et \mathcal{F} sont des fonctions à valeurs réelles auxquelles on peut associer plusieurs propriétés additionnelles, par exemple, la non-décroissance ou la croissance dans tous ses arguments.

Définition 2.3.1 Une mesure multidimensionnelle de la pauvreté se définit comme une fonction non-constante $\mathbf{P} : Z \times X \rightarrow \mathbb{R}$ telle que :

– Type d'agrégation I :

$$\begin{aligned} \mathbf{P}(\mathbf{z}, X) &= G(\mathbf{f}^1(\mathbf{z}, X), \mathbf{f}^2(\mathbf{z}, X), \dots, \mathbf{f}^m(\mathbf{z}, X)) \\ &= G(\pi^1, \pi^2, \dots, \pi^m) \end{aligned} \quad (2.4)$$

– Type d'agrégation II :

$$\begin{aligned} \mathbf{P}(\mathbf{z}, X) &= F(\mathbf{f}_1(\mathbf{z}, X), \mathbf{f}_2(\mathbf{z}, X), \dots, \mathbf{f}_n(\mathbf{z}, X)) \\ &= F(\pi_1, \pi_2, \dots, \pi_n) \end{aligned} \quad (2.5)$$

$G(\pi_1, \pi_2, \dots, \pi_n)$ est une application de $\mathbb{R}_+^n \rightarrow \mathbb{R}_+$ qui définit une procédure d'agrégation des indicateurs cardinal de pauvreté π^j en une mesure multidimensionnelle de la pauvreté $\mathbf{P}(\mathbf{z}, X)$. Tandis que, $F(\pi_1, \pi_2, \dots, \pi_n)$ est une application de $\mathbb{R}_+^n \rightarrow \mathbb{R}_+$ qui définit une procédure d'agrégation des indicateurs composites de la pauvreté π_i en une mesure multidimensionnelle de la pauvreté $\mathbf{P}(\mathbf{z}, X)$. Pour tout X et $\mathbf{z} \in Z$, $\mathbf{P}(\mathbf{z}, X)$ représente le niveau de pauvreté associé à la matrice des réponses X selon un vecteur des seuils de pauvreté \mathbf{z} . En pratique, la forme explicite de $\mathbf{z} \in Z$, et les fonctions $G(\cdot, \cdot)$, $F(\cdot, \cdot)$, $\mathbf{f}^j(\cdot, \cdot)$ et $\mathbf{f}_i(\cdot, \cdot)$ sont généralement choisies (ou construites) telles que $\mathbf{P}(\mathbf{z}, X)$ satisfasse certaines propriétés ou certains axiomes.

Soit n est un entier naturel supérieur ou égal à 2, on désigne par \mathfrak{D}_n le groupe des permutations de $i \in \{1, 2, \dots, n\}$ et par $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ la base canonique de \mathbb{R}^n . Pour tout couple (i, j) d'entiers naturels, on note δ_{ij} le symbole de Kronecker ($\delta_{i,i} = 1$ et $\delta_{i,j} = 0$ pour $i \neq j$).

Définition 2.3.2 Soit $\sigma \in \mathfrak{D}_n$, on appelle matrice de permutation associée à σ , la matrice de passage Π_σ de la base canonique de \mathbb{R}^n à la base $\mathcal{B} = \{e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}\}$. On a donc, si Π_σ est une matrice de permutation, $\Pi_\sigma e_j = e_{\sigma(j)}$ pour tout entier j ; ce qui revient à dire que :

$$\Pi_\sigma = (\delta_{i,\sigma(j)})_{1 \leq i, j \leq n} \quad (2.6)$$

Définition 2.3.3 On dit qu'une matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{R})$ est stochastique si elle est positive et :

$$\forall i \in \{1, 2, \dots, n\}, \quad \sum_{j=1}^n a_{ij} = 1 \quad (2.7)$$

où $\mathcal{M}_n(\mathbb{R})$ représente l'ensemble des matrices carrées d'ordre n .

Exemple 2.3.1 Une matrice de permutation est stochastique mais l'inverse n'est pas vrai.

Définition 2.3.4 On appelle matrice bistochastique (ou doublement stochastique) une matrice stochastique A telle que A^T soit aussi stochastique.

Exemple 2.3.2 Les matrices de permutation sont des matrices bistochastiques mais l'inverse n'est pas vrai.

2.3.2 Les approches traditionnelles

Selon Sen [1979b], le problème de mesure de la pauvreté peut être scindé en deux phases : (i) l'identification des pauvres parmi la population totale (qui sont les pauvres ?) et (ii) l'agrégation (comment combiner les caractéristiques des différentes personnes en une mesure agrégée ?). Dans l'approche traditionnelle, le problème (i) a été résolu par une approche basée sur le revenu (ou la consommation), laquelle nécessite la spécification d'un revenu de subsistance limite (ou ligne de pauvreté). Ainsi, une personne sera identifiée comme "pauvre" lorsque son revenu (ou sa consommation) se trouve en dessous de ce niveau de coupure. De nombreuses procédures ont été proposées par la littérature pour le traitement du problème (ii), lesquelles sont basées sur la construction d'un indice de la pauvreté utilisant les informations disponibles sur les personnes. Nous présentons dans cette section une revue des mesures de la pauvreté les plus populaires.

2.3.2.1 L'indice de pauvreté de Foster, Greer et Thorbecke

L'approche traditionnelle, appelée aussi approche monétaire, calcule le revenu (ou la consommation) des individus comme indicateur du bien-être. Premièrement, on définit un indicateur de bien-être ζ_i du i ème ménage et le seuil de pauvreté z tels que, $\zeta_i = \sum_{j=1}^m \mu_j x_i^j$, où μ_j représente le prix du marché du bien j , et $z = \sum_{j=1}^m \mu_j z^j$. La fonction d'agrégation (ou indice de pauvreté de Foster-Greer-Thorbecke) est

donnée par :

$$\mathbf{P}(\mathbf{z}, X) = FGT_\gamma = F(\mathbf{f}_1(\mathbf{z}, X), \mathbf{f}_2(\mathbf{z}, X), \dots, \mathbf{f}_n(\mathbf{z}, X)) \quad (2.8)$$

$$= \frac{1}{n} \sum_{i=1}^n (\mathbf{f}_i(\mathbf{z}, X))^\gamma \mathbf{I}(\mathbf{f}_i(\mathbf{z}, X)) > 0 \quad (2.9)$$

$$= \frac{1}{n} \sum_{i=1}^n \left[\max \left\{ \frac{\mathbf{z} - \zeta_i}{\mathbf{z}}; 0 \right\} \right]^\gamma \mathbf{I}(\mathbf{z} > \mathbf{f}_i(\mathbf{z}, X)) \quad (2.10)$$

où $\mathbf{f}_i(\mathbf{z}, X) = \max \left[\frac{\mathbf{z} - \zeta_i}{\mathbf{z}}; 0 \right]$ est la fonction de pauvreté individuelle, \mathbf{I} une fonction indicateur et γ le paramètre de sensibilité de l'indice de distribution parmi les personnes pauvres. En particulier, si $\gamma = 0$, FGT_0 représente "l'incidence de la pauvreté" (aussi appelée "*indice en nombre d'habitants*") c'est-à-dire la proportion de la population dont le revenu ou la consommation se situe en dessous du seuil de pauvreté. Cette indice s'interprète comme la part de la population qui ne peut pas se permettre d'acheter le panier de produits correspondant au minimum vital. Si $\gamma = 1$, FGT_1 représente la "profondeur de la pauvreté" (aussi appelé "*écart de pauvreté*") c'est-à-dire la distance à laquelle les ménages se trouvent du seuil de pauvreté. Cet indice s'interprète comme le déficit collectif moyen de revenu ou de consommation par rapport au seuil de pauvreté pour l'ensemble de la population. Et si $\gamma = 2$, FGT_2 représente la "sévérité de la pauvreté" (aussi appelé "*écart de pauvreté au carré*") qui tient compte non seulement de la distance séparant les pauvres du seuil de pauvreté (l'écart de pauvreté), mais aussi de l'inégalité entre les pauvres. Elle attribue une pondération plus importante aux ménages situés à plus grande distance de la ligne de pauvreté. Notons que ces indices appartiennent à la famille des mesures de la pauvreté appelée *indices de Foster-Greer-Thorbecke* [voir Foster *et al.*, 1984].

2.3.2.2 L'indice de pauvreté de Sen

Sen [1976] a utilisé une approche axiomatique pour dériver un autre indice de pauvreté. L'indice de Sen, noté $\mathcal{S}(\mathbf{z}, X)$, est donné par l'équation (2.11) :

$$\mathbf{P}(\mathbf{z}, X) = \mathcal{S}(\mathbf{z}, X) = \left(\frac{2}{(q+1)n\mathbf{z}} \right) \sum_{i=1}^q (\mathbf{z} - \zeta_i)(q+1-i) \quad (2.11)$$

L'équation (2.11) peut aussi s'écrire comme suite :

$$\mathbf{P}(\mathbf{z}, X) = \mathcal{S}(\mathbf{z}, X) = FGT_0 \left(1 - (1 - FGT_1) \left(1 - \mathbf{G} \left(\frac{q}{q+1} \right) \right) \right) \quad (2.12)$$

où \mathbf{G} se réfère au coefficient de Gini de la distribution de Lorenz des revenus ou consommation des pauvres [voir Gini, 1913; Theil, 1967] et q est le nombre de

pauvres. Notons que lorsque q est suffisamment grand, on peut aisément montrer que l'équation (2.12) est équivalente à l'équation (2.13) suivante :

$$\mathbf{P}(\mathbf{z}, X) = \mathcal{S}(\mathbf{z}, X) = FGT_0(1 - (1 - FGT_1)(1 - \mathbf{G})) \quad (2.13)$$

Contrairement à FGT_0 (la proportion de la population dont le revenu ou la consommation se situe en dessous du seuil de pauvreté), $\mathcal{S}(\mathbf{z}, X)$ est sensible à l'étendue de la distance des revenus des pauvres du seuil de pauvreté. Contrairement à FGT_1 (l'écart entre le seuil de pauvreté et le revenu moyenne des pauvres, exprimé comme une proportion du seuil de pauvreté), $\mathcal{S}(\mathbf{z}, X)$ est sensible au nombre de pauvres en dessous du seuil de pauvreté. De plus, $\mathcal{S}(\mathbf{z}, X)$ n'est pas insensible à la redistribution des revenus [voir Sen, 1976, pour plus de détails]. De là, l'indice de Sen $\mathcal{S}(\mathbf{z}, X)$ contraste fortement avec les deux mesures brutes de la pauvreté (FGT_0 et FGT_1) utilisées dans la littérature statistique et dans les débats politiques sur la pauvreté.

2.3.2.3 L'indice de pauvreté de Watts

L'indice de pauvreté de Watts, noté $\mathcal{W}(\mathbf{z}, X)$, se définit comme suit :

$$\mathbf{P}(\mathbf{z}, X) = \mathcal{W}(\mathbf{z}, X) = \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{\mathbf{z}}{\zeta_i} \right) \mathbf{I}(\zeta_i \leq \mathbf{z}) \quad (2.14)$$

Dans ce cas, ζ_i est considéré comme le revenu (positif) de l'individu i . L'indice de Watts [1968] a la particularité de satisfaire l'axiome de focalisation, l'axiome de monotonie, l'axiome de transfert et l'axiome de décomposabilité (voir section 2.3.4.2) que les indices FGT_0 et FGT_1 ne satisfont pas. Le logarithme joue un rôle important au niveau de la sensibilité dans le sens où il garantit qu'un transfert pur de revenu des plus pauvres des pauvres vers ceux qui sont mieux nantis modifiera l'indice de Watts. Toutefois, cet indice n'est pas particulièrement intuitif et est très rarement utilisé en pratique sur le terrain.

2.3.2.4 Discussion

Bien que les indices de pauvreté analysés dans cette section soient les plus populaires avec l'avantage d'être simple à construire et de refléter ce qu'ils sont censés capturer, ils souffrent néanmoins de quelques inconvénients. FGT_0 est "complètement insensible à la redistribution du revenu parmi les pauvres. Un transfert pur de revenu des plus pauvres des pauvres vers ceux qui sont mieux nantis gardera l'indice FGT_0 inchangé, ou elle va diminuer - ce qui correspond à un faux résultat." [see Sen, 1976]. FGT_1 ne capture pas les différences dans la sévérité de la pauvreté entre les pauvres et ignore également l'inégalité entre les pauvres.

Ainsi, FGT_0 ne satisfait pas l'axiome de monotonie et l'axiome de transfert, tandis que FGT_1 ne satisfait pas uniquement l'axiome de transfert. FGT_2 est très difficile à comprendre et à interpréter. L'indice de pauvreté de Sen dépend du coefficient Gini de Gini et partage alors ses principaux inconvénients. En effet, l'indice de Gini (et par conséquent l'indice de Sen) ne peut pas être utilisé pour décomposer la pauvreté en distributions de sous-groupes différents. L'indice de Watts est distributionnellement sensible en vertu de l'utilisation du logarithme.

Les approches traditionnelles reposent sur le revenu ou les dépenses de consommation par habitant pour calculer des indices différents. Par conséquent, elles explorent uniquement la dimension monétaire de la pauvreté en omettant le fait que la pauvreté est multidimensionnelle. Alors nous avons aussi à prendre en compte les informations non-monétaires lors de la conceptualisation d'une mesure de la pauvreté. Nous devons discuter de ce que nous entendons par multidimensionnelle et en tenir compte pendant le mesurage. Dans la pratique, les mesures monétaires sont pour la plupart déconnectées des valeurs, de la réalité et de la façon dont les gens se perçoivent eux-mêmes. On note ici un risque élevé de biais ethnocentrique dû au fait que la liste des besoins de base tend à refléter les points de vue d'un politicien ou d'un analyste, et non ceux de la population elle-même.

Notre vision de la pauvreté est basée sur la combinaison de trois questions :

- les données empiriques sur les conditions de vie des ménages qui doivent être considérées pour comprendre à la fois ce que la situation actuelle 'est' et pour prévoir les moyens possibles d'y remédier ;
- considérant la pauvreté, nous devons prendre en compte à la fois la perception subjective de ce que signifie "être pauvre" (capturé par le concept d'*être limité* dans l'atteinte de ses aspirations) et la perception religieuse, philosophique et idéologique, ainsi que les moyens par lesquels les hommes prennent conscience de ce fléau et essayent de le combattre ;
- le mesurage de la pauvreté doit être considéré comme un instrument pour concevoir, élaborer et évaluer les politiques de réduction de la pauvreté et non comme une activité pour elle-même. Notre revendication est qu'il n'existe aucun autre objectif du mesurage de la pauvreté que les politiques et les décisions qui en découleront.

2.3.3 Les approches multidimensionnelles traditionnelles

Les approches décrites dans cette section sont censées mesurer la déprivation totale au sein d'une société en terme de chaque attribut séparément, puis d'agréger les différents indices en un indice unique.

2.3.3.1 Les approches d'évaluation de la possession sur chaque attribut

Les approches d'évaluation uniattribut concernent l'évaluation de la possession ou de la déprivation de chaque indicateur élémentaire. Ces approches nous permettent d'évaluer dans quelle mesure les individus sont déprivés ou non sur chaque dimension ou item pris individuellement.

a) Les approches basées sur le seuil (cut-off)

Townsend [1979, 1987] fût le premier à introduire le concept de “déprivation” suite à la critique faite concernant le *paradigme monétaire*. La méthode “cut-off” consiste à évaluer la déprivation sur chaque dimension par un classement binaire du type de déprivation ou non. Ainsi, un individu (ou un ménage) sera dit pauvre par rapport à un attribut lorsque la réalisation de cet attribut est en dessous d'un seuil objectif ou d'une norme sociale.

Formellement, on construit une fonction croissante $f : Z \times X_j \rightarrow \{0; 1\}$ sur chaque attribut laquelle est alors utilisée pour affecter les individus à l'ensemble des “déprivés” D_j sur l'attribut X_j . La fonction $f(z^j, x_i^j)$ est donnée par l'équation (2.15) suivante :

$$f(z^j, x_i^j) = \begin{cases} 1 & \text{si } x_i^j \geq z^j \Rightarrow \mathbf{x}_i \in D_j. \\ 0 & \text{si } x_i^j < z^j \Rightarrow \mathbf{x}_i \in \overline{D}_j. \end{cases} \quad (2.15)$$

où $D_j \subseteq \Omega$ et $\overline{D}_j = \{\mathbf{x}_i : \mathbf{x}_i \in \Omega \wedge \mathbf{x}_i \notin D_j\}$. Pour les applications de cette méthode, le lecteur intéressé pourra consulter l'ouvrage de Townsend [1979] et celui de Mack et Lansley [1985]. Cette méthode est généralement utilisée pour traiter des problèmes de pauvreté qui font appel aux items dichotomiques.

Cette méthode présente de nombreux inconvénients, notamment la détermination d'une procédure consensuelle pour attribuer des valeurs dichotomiques à certaines modalités. Par exemple, pour les deux modalités telles que “ne pas consommer de l'eau potable” ou “ne pas pouvoir se nourrir convenablement”, il est communément accepté qu'elles constituent des désavantages qui sont difficiles à juger par des valeurs numériques. À cela vient s'ajouter le problème de détermination des seuils z^j définis comme le minimum objectif ou social sur l'attribut X_j . D'autre part, cette méthode ne nous renseigne pas sur la qualité et quantité de l'item mais uniquement sur la présence ou l'absence de l'item. De plus, l'évaluation de la déprivation dans cette méthode fait l'hypothèse qu'il existe seulement deux situations possibles (soit pauvre, soit non pauvre) sans prendre en compte la nature vague du concept de pauvreté.

b) Les approches basées sur la théorie des ensembles flous

2.3.3.2 * La théorie des ensembles flous

La théorie des ensembles flous a été introduite par Zadeh [1965]. Elle repose sur un principe de gradualité qui étend la logique classique bivalente caractérisée par l'opposition stricte "vrai-faux". Selon Zadeh [1965], les classes d'objets rencontrées dans le monde physique n'ont pas de critères d'appartenance précisément définis mais sont plutôt caractérisées par un continuum de degré d'appartenance. L'absence de critères précis d'appartenance fait que ces ensembles permettent d'appréhender les concepts vagues tels qu'*être jeune* ou *être pauvre*. Du fait qu'il n'existe pas toujours une *ligne de démarcation absolue* entre *pauvres* et *non pauvres*, la théorie des ensembles flous permet alors d'assigner une valeur communément appelée *degré de vérité* à une proposition telle que "*être non pauvre*" ou "*être pauvre*". Les paradoxes Sorites tels que le paradoxe du barbu, le paradoxe du riche, le paradoxe du tas sont souvent utilisés pour illustrer le concept vague. Par exemple, le paradoxe du barbu intervient dans ce qu'il n'est pas toujours vrai qu'enlever un simple poil de barbe à quelqu'un qui en possède beaucoup pour qu'il devienne non barbu. Mais, si on lui enlève beaucoup de poils, un par un, l'individu deviendra non barbu sans que l'on puisse dire à quel moment exact se fait le passage à l'état de non barbu. En effet, il n'existe pas une ligne de démarcation claire entre les situations de présence ou d'absence de barbe.

2.3.3.3 * La théorie de la supervaluation

La théorie des ensemble flous a été adoptée par certains auteurs [Chiappero, 1994; Cheli et Lemmi, 1995; Qizilbash, 2000, 2003; Cerioli et Zani, 1990; Appiah-Kubi *et al.*, 2007] pour aborder la question de mesure de la pauvreté. Qizilbash [2000, 2003] a proposé deux méthodes pour appréhender le caractère vague des concepts :

La première méthode consiste à appréhender l'imprécision (*vagueness*) comme le résultat de notre *ignorance* quant à la position des contours d'un concept. En effet, l'impossibilité de tracer une ligne de démarcation claire entre la présence ou l'absence d'un attribut provient du fait qu'on ne sait pas où la situer et non parce qu'elle n'existe pas. Ravallion [1996] avait déjà été confronté à un dilemme semblable et avait tenté de le résoudre en proposant l'approche par des seuils multiples. Toutefois, cette vision de l'imprécision n'est pas satisfaisante pour l'analyse multidimensionnelle de la pauvreté car elle ne permet pas de résoudre les problèmes de *l'imprécision verticale* et de *l'imprécision horizontale*.

La seconde méthode est la *théorie de la supervaluation* [Qizilbash, 2000, 2003] qui suggère que lorsqu'un concept est vague, la vérité d'un état qui implique le concept peut dépendre de la manière dont le concept peut être rendu plus précis.

Dès lors, dans le cadre de la théorie de la “supervaluation”, quelqu’un pourra sans hésitation être considéré comme *pauvre* s’il est en-dessous de *tous* les seuils critiques admissibles sur une dimension qui fait partie de *toutes* les spécifications admissibles du concept. Si on constate qu’il est *super-vrai* que la personne est pauvre alors elle est dite “*core poor*”. Si un individu ne souffre d’aucune insuffisance dans aucune des dimensions d’aucune liste sur aucun niveau critique alors il est *super-vrai* que cette personne n’est pas pauvre. On dira alors que cet individu est “*riche*” ou “*core non poor*”. Enfin, lorsqu’un individu n’est ni “*core poor*” et ni “*core non poor*”, elle est aux marges de la pauvreté. Le lecteur intéressé par cette approche pourra consulter Qizilbash [2003]; Fusco [2005].

2.3.3.4 ★ Formulation

Soit \mathcal{P} un sous-ensemble de pauvres et $\mathcal{P} \subset \Omega$. Un ensemble flou $\tilde{\mathcal{P}}$ de \mathcal{P} est défini par l’équation (2.16) suivante :

$$\tilde{\mathcal{P}} = \{(\mathbf{x}_i, \mathbf{f}_i(\mathbf{z}, X))\}, \quad \mathbf{x}_i \in \mathcal{P}. \quad (2.16)$$

où $\mathbf{f}_i(\mathbf{z}, X)$ est le degré d’appartenance de l’individu \mathbf{x}_i au sous-ensemble \mathcal{P} . Alors, nous avons :

$$\begin{cases} \mathbf{f}_i(\mathbf{z}, X) = 0 & \implies \mathbf{x}_i \in \Omega - \mathcal{P}, & \mathbf{x}_i \text{ n'appartient pas au sous-ensemble } \mathcal{P}; \\ 0 < \mathbf{f}_i(\mathbf{z}, X) < 1 & \implies \mathbf{x}_i \in \tilde{\mathcal{P}}, & \mathbf{x}_i \text{ appartient partiellement au sous-ensemble } \mathcal{P}; \\ \mathbf{f}_i(\mathbf{z}, X) = 1 & \implies \mathbf{x}_i \in \mathcal{P}, & \mathbf{x}_i \text{ appartient complètement au sous-ensemble } \mathcal{P}. \end{cases}$$

où $\Omega - \mathcal{P} = \{\mathbf{x}_i : \mathbf{x}_i \in \Omega \wedge \mathbf{x}_i \notin \mathcal{P}\}$ et $\tilde{\mathcal{P}}$ l’appartenance partielle. En pratique, $\mathbf{f}_i(\mathbf{z}, X)$, donnée par l’équation (2.3), peut-être définie par :

$$\mathbf{f}_i(\mathbf{z}, X) = \left[\sum_{j=1}^m \lambda_j [f(z^j, x_i^j)]^\gamma \right]^{1/\gamma} \quad (2.17)$$

où γ représente le paramètre de sensibilité; λ_j , le système de pondération associé à chaque attribut j . De là, $\mathbf{f}_i(\mathbf{z}, X)$ se réfère à l’indice agrégé de degrés d’appartenance $f(z^j, x_i^j)$ associé à l’indicateur de déprivation X_j . Dans la suite, nous présentons deux instances de cette approche.

2.3.3.5 Quelques approches basées sur la théorie des ensembles floues

★ L’approche totalement floue

Cette approche a été introduit par Cerioli et Zani [1990] pour appliquer le concept de la théorie des ensembles floues au mesurage de la pauvreté. Cerioli et Zani [1990] définit la fonction d’appartenance $g : \Omega \longrightarrow \{0, 1\}$ selon le fait que les variables sont dichotomiques, catégorielles ou continues.

Dans le cas des *variables dichotomiques* (variable à deux modalités), la fonction d'appartenance g est donnée par l'équation (2.18) :

$$g(x_i^j) = \begin{cases} 1 & \text{si } x_i^j = 1 \Rightarrow \mathbf{x}_i \in D_j. \\ 0 & \text{si } x_i^j = 0 \Rightarrow \mathbf{x}_i \in \overline{D}_j. \end{cases} \quad (2.18)$$

où $D_j \subseteq \Omega$ et $\overline{D}_j = \{\mathbf{x}_i : \mathbf{x}_i \in \Omega \wedge \mathbf{x}_i \notin D_j\}$.

Dans le cas des *variables catégorielles* (variables avec k modalités ordonnées, $k > 2$), la fonction d'appartenance g est donnée par l'équation (2.19) :

$$g(x_i^j) = \begin{cases} 1 & \text{if } x_i^j \geq x_{\max}^j \Rightarrow \mathbf{x}_i \in D_j. \\ \frac{x_i^j - x_{\min}^j}{x_{\max}^j - x_{\min}^j} & \text{if } x_i^j \in [x_{\min}^j; x_{\max}^j[\Rightarrow \mathbf{x}_i \tilde{\in} D_j. \\ 0 & \text{if } x_i^j < x_{\min}^j \Rightarrow \mathbf{x}_i \in \overline{D}_j. \end{cases} \quad (2.19)$$

où x_{\min}^j est la plus petite valeur associée à l'évaluation d'un individu de référence \mathbf{x}_{\inf} et x_{\max}^j est la plus grande valeur associée à l'évaluation de l'individu de référence \mathbf{x}_{\sup} , sur l'attribut $j \in \{1, \dots, m\}$. La notation $\tilde{\in}$ signifie que l'individu est '*partiellement déprivé*' sur le j ème attribut. Notons que si $x_{\min}^j = x_{\max}^j$, l'équation (2.19) devient équivalent à l'équation (2.15).

Dans le cas des *variables continues* (telle que le revenu ou les dépenses de la consommation), Cerioli et Zani [1990] définissent deux références x_{\min}^j et x_{\max}^j telles que si la valeur x_i^j prise par l'indicateur continue pour un individu donné est plus petite que x_{\min}^j , cet individu sera incontestablement considéré comme pauvre sur le j ème attribut. Tandis que si elle est plus grande que x_{\max}^j , il sera considéré comme certainement pauvre sur le j ème attribut. La fonction d'appartenance g est alors donnée par l'équation (2.20) :

$$g(x_i^j) = \begin{cases} 0 & \text{if } x_i^j \geq x_{\max}^j \Rightarrow \mathbf{x}_i \in \overline{D}_j. \\ \frac{x_{\max}^j - x_i^j}{x_{\max}^j - x_{\min}^j} & \text{if } x_i^j \in [x_{\min}^j; x_{\max}^j[\Rightarrow \mathbf{x}_i \tilde{\in} D_j. \\ 1 & \text{if } 0 \leq x_i^j < x_{\min}^j \Rightarrow \mathbf{x}_i \in D_j. \end{cases} \quad (2.20)$$

La fonction d'appartenance g ainsi obtenue est une fonction croissante et linéaire évoquant le fait que le risque de pauvreté augmente linéairement entre x_{\min}^j et x_{\max}^j . L'une des difficultés de cette méthode concerne la détermination de façon pertinente des valeurs des seuils supérieur et inférieur. Les auteurs proposent de fixer x_{\min}^j au seuil de subsistance et x_{\max}^j au niveau de la moyenne du revenu par tête. On retrouve dans cette proposition intéressante l'idée de l'utilisation conjointe d'un seuil absolu et d'un seuil relatif [voir Ravallion, 1996, 1998, 2008, pour plus de détails]. Ainsi, les personnes atteintes de déprivation sévère seront considérées comme complètement pauvres tandis que celles ayant un revenu supérieur à la moyenne ne le seront pas du tout.

★ L'approche relative totalement floue

Cette approche, due à Cheli et Lemmi [1995] est une amélioration de l'approche originelle proposée par Cerioli et Zani [1990] en ce qu'elle permet d'éviter la spécification des seuils critiques inférieur et supérieur (totalement floue) et l'utilisation d'une relation linéaire reposant sur l'hypothèse discutable d'une équidistance entre les modalités. En plus, le degré de déprivation de chaque individu sur une variable donnée dépendra de sa place dans la distribution de la variable.

Dans le cas des *variables dichotomiques*, la fonction d'appartenance g est définie de la même manière que l'équation (2.18).

Dans le cas des *variables catégorielles* avec k modalités, l'hypothèse d'équidistance entre les modalités présente chez Cerioli et Zani [1990] est relâchée. La fonction d'appartenance g se définit comme suite :

$$g(x_i^j) = \begin{cases} 0 & \text{si } x_i^j = \beta_1^j. \\ g(\beta_{q-1}^j) + \frac{\mathbf{F}_j(\beta_q^j) - \mathbf{F}_j(\beta_{q-1}^j)}{1 - \mathbf{F}_j(\beta_1^j)} & \text{si } x_i^j = \beta_q^j \end{cases} \quad (2.21)$$

où β_q^j désigne les $q \in \{1, \dots, k\}$ modalités que X_j peut prendre. La modalité β_1^j a le plus faible risque de pauvreté et la modalité β_q^j , le plus grand risque de pauvreté. \mathbf{F}_j est la fonction de distribution cumulative de X_j .

Dans le cas de variables continues, la fonction d'appartenance g ainsi obtenue est une fonction croissante mais non linéaire. On peut ainsi dépasser la linéarité de la fonction d'appartenance.

Cette approche est moins arbitraire car on n'a plus besoin de définir les seuils et elle est cohérente avec une approche relative de la pauvreté. En plus, les degrés d'appartenance relatifs à cette distribution fictive ne sont plus équidistants et dépendent uniquement de la distribution de l'attribut dans la population.

Remarque 2.3.1 *Cette spécification est attractive mais présente une limite dont souffre toute approche relative. Les degrés d'appartenance ne sont déterminés qu'en fonction de la position relative de chaque individu par rapport aux autres individus ; ainsi, seule l'information sur le pourcentage des individus ayant un revenu supérieur est utilisée, et pas celle concernant les écarts de pauvreté, ni les niveaux de revenus. Dès lors, on pourrait aboutir au même ensemble de degré d'appartenance pour deux distributions différentes (l'une ayant des revenus concentrés et l'autre plus dispersés).*

★ Discussion

Les limites des approches basées sur la théorie des ensembles ont été amplement discutées par Qizilbash [2000, 2003]; Fusco [2005]. De manière générale, les

insuffisances de ces approches apparaissent lorsque les applications multidimensionnelles sont basées sur un indice composite conçu comme une moyenne agrégée. En effet, les implications de cette méthode sont les suivantes : soit un *individu est pauvre* avec un degré de vérité égal à 1 indiquant le fait qu'il est *pauvre dans toutes les dimensions*, soit un *individu est non-pauvre* avec un degré de vérité égal à 0 indiquant le fait qu'il est *non pauvre sur toutes les dimensions*, ou alors l'individu est pauvre de manière non ambiguë sur certaines dimensions et non pauvre de manière non ambiguë sur d'autres (c'est-à-dire qu'il est pauvre à un certain degré). Ce raisonnement est incompatible avec une vision selon laquelle un individu est pauvre si un de ses besoins essentiels n'est pas satisfait. Si on se situe dans une approche absolue où les dimensions sont toutes indispensables et font toutes partie d'un noyau de pauvreté, cela n'est pas satisfaisant.

D'autre part, l'approche par les ensembles flous distingue de manière rigide deux catégories (l'ensemble des pauvres et celui des non pauvres). Mais, il devient difficile de déterminer et de justifier l'existence des autres catégories. Toutefois, malgré les aspects théoriques séduisants de la théorie, la méthode des ensembles flous a du mal à s'inscrire sur le plan international et son application semble ne pas répondre parfaitement aux problèmes liés au contexte de la mesure de la pauvreté. Les raisons à cela sont certainement à chercher dans les difficultés d'interprétation de ce que nous avons exposé plus haut. A ces difficultés, viennent se greffer le problème de normalisation d'échelles des différents indicateurs et la question d'agrégation de ces différents indicateurs de déprivation pour obtenir les indices agrégés.

2.3.3.6 Les approches basées sur les indicateurs agrégés

Le principe des méthodes d'agrégation est de combiner et de résumer plusieurs grandeurs numériques en un indice appelé *indice composite global*, de telle sorte que l'indice prenne en considération d'une manière prescrite, toutes les valeurs individuelles.

* L'indice de pauvreté humaine (*IPH*)

L'*IPH* [PNUD, 1990–2006] est basé sur trois indices IPH_1 , IPH_2 et IPH_3 , respectivement les *déficiences en termes de santé* (proportion de mortalité avant l'âge de 40 ans), les *déficiences en termes d'éducation* (analphabétisme) et les *déficiences en termes de conditions de vie* (accès à l'eau potable, accès aux soins de santé et nutrition infantile). L'indice composite global *IPH* a été formulé par Armand et Sen [1997] comme suit :

$$IPH = \left(\lambda_1 IPH_1^\gamma + \lambda_2 IPH_2^\gamma + \lambda_3 IPH_3^\gamma \right)^{1/\gamma} \quad (2.22)$$

avec $\lambda_1 + \lambda_2 + \lambda_3 = 1$ et $\gamma \geq 1$ est un paramètre. Pour $\gamma = 1$, les composantes de *IPH* sont parfaitement substituables et ce dernier n'est que leurs espérances mathématiques. Par contre, lorsque $\gamma \rightarrow \infty$, cet indice tendra vers la valeur maximale de ses trois composantes, c'est-à-dire que plus de poids est accordé à la composante où le déficit est plus élevé ($\max(IPH_1, IPH_2, IPH_3)$). Dans ce cas, l'*IPH* ne baisse que si sa composante la plus élevée diminue.

★ L'approche complémentaire

Brandolini et D'Alessio [1998] explorent la possibilité d'une analyse multidimensionnelle de la déprivation et de l'inégalité en exploitant l'idée de l'approche par les capacités de Sen [voir Sen, 1985]. Ils examinent alors une stratégie différentes pour mesurer les fonctionnements et les capacités. Dans ce contexte, ils définissent l'approche complémentaire (*supplementary strategy*, en anglais), laquelle vise à compléter l'information dérivée de la distribution des ressources (mesurée par le revenu ou la consommation) par les indicateurs du niveau de vie. Dans cette stratégie, les fonctionnements sont examinés variable par variable de manière à décrire la caractéristique de chaque dimension et d'étudier leur corrélation.

L'avantage de cette technique est sa simplicité. En fait, elle n'impose aucune contrainte particulière sur la structure des données sur la pauvreté et les prérequis du mesurage sont moins contraignants. L'absence de synthèse et la difficulté de donner une image unitaire bien définie sont les principaux inconvénients de cette méthode. En effet, la pluralité des indicateurs élémentaires conflictuels existant simultanément au sein d'une même modélisation peut conduire à une difficulté d'analyse et de synthèse. En plus, il est fort probable d'obtenir uniquement un ordre partiel lors des comparaisons des observations entre elles. Cela constitue un argument majeur en faveur de la construction d'indices agrégés qui constituent un moyen de synthétiser l'information

★ L'indice composite global

L'approche par les indices agrégés consiste à construire un indicateur composite global $f_i(\mathbf{z}, X)$ pour chaque ménage x_i . Étant donné une matrice réponses X , le problème posé est alors de déterminer une fonction d'agrégation \mathcal{F} définie de \mathbb{R}^m dans \mathbb{R} tel que :

$$f_i(\mathbf{z}, X) = \mathcal{F}(f(z^1, x_i^1), \dots, f(z^m, x_i^m)) \quad (2.23)$$

La fonction \mathcal{F} (voir équation(2.3)) peut être définie de différentes manières selon les auteurs [Chiappero, 1994; Bourguignon et Chakravarty, 2003] et les propriétés que l'on veut associer à cette fonction. La fonction d'agrégation la plus répandue est la *moyenne pondérée d'ordre γ* telle que donnée par l'équation (2.17),

où $f_i(\mathbf{z}, X)$ représente l'indice agrégé pour l'individu \mathbf{x}_i des degrés normalisés $f(z^j, x_i^j)$ associés aux différents indicateurs élémentaires de déprivation; λ_j le poids associé à chaque attribut j . Généralement, on a $\lambda_j \geq 0$ et $\sum_{j=1}^m \lambda_j = 1$; γ est un paramètre qui détermine le niveau de substitution entre les attributs. γ fait également référence au type de moyenne. Par exemple, $\gamma = 0$ implique la moyenne géométrique; $\gamma = -1$, la moyenne harmonique et $\gamma = 1$, la moyenne arithmétique. Le problème de choix d'un système de pondération approprié pour différents indicateurs de bien-être a été abordé par Ram [1982] qui a suggéré d'utiliser les données pour déterminer le poids optimal associé à chaque attribut. Ce type d'opérateur d'agrégation comporte plusieurs inconvénients dont les plus flagrants sont la commensurabilité des dimensions et leur caractère compensatoire.

Ce type d'opérateur d'agrégation implique la commensurabilité des variables et la compensation/substituabilité [voir Fusco, 2005; Atkinson, 2003]. Considérons la variable X_j : salaire mensuelle. Admettons que ce soit une graduation et que la proposition suivante soit jugée réaliste : "toutes choses égales par ailleurs, un accroissement du salaire mensuel de μ francs CFA⁵ procure un supplément de satisfaction μ fois supérieur à celui que procure un accroissement de 100 francs CFA". Une graduation X_j est conforme à l'exigence de commensurabilité des écarts de préférences selon son axe de signification j (avec $j = 1 \dots m$), si pour tout $(\mathbf{x}_i, \mathbf{x}_k, \mathbf{x}_u, \mathbf{x}_v) \in [\Omega^4]_j$, l'égalité 2.24 apparaît réaliste pour définir la restriction de la fonction ϕ à cet ensemble.

$$\phi\left(\frac{\mathbf{x}_i \ominus \mathbf{x}_k}{\mathbf{x}_u \ominus \mathbf{x}_v}\right) = \frac{x_i^j - x_k^j}{x_u^j - x_v^j} \quad (2.24)$$

La théorie du mesurage [see Roberts, 1979; Krantz *et al.*, 1971] établit que les seuls codages qui transforment une graduation x_i^j conforme à l'exigence de commensurabilité des écarts de préférence en une autre graduation $\chi(x_i^j)$ conforme à cette même exigence sont des transformations affines données par l'équation 2.25 :

$$\chi(x_i^j) = \alpha x_i^j + \beta \quad (2.25)$$

où $\alpha \in \mathbb{R}_+$ et $\beta \in \mathbb{R}$. Ce point de vue, qui vise à garantir la signifiante de la mesure, est souvent ignoré dans la plupart des approches multidimensionnelles de mesure de la pauvreté basées sur l'idée des comparaisons deux à deux.

D'autre part, le niveau de compensation/substituabilité dépend de la valeur γ et de l'ensemble $\{\lambda_j : j = 1 \dots m\}$. Par exemple, si $\gamma = 1$ et $\lambda_j = \frac{1}{m}$, $\forall j \in \{1, \dots, m\}$ alors $f_i(\mathbf{z}, X) = \frac{1}{m} \sum_{j=1}^m f(z^j, x_i^j)$. Ainsi, la fonction est additive et la compensation est totale (substituabilité infinie) : l'augmentation d'une unité dans

5. The CFA franc is the name of two currencies (the West African CFA franc and the Central African CFA franc) used in Africa which are guaranteed by the French treasury. 1 euro = 655.957 CFA francs.

une dimension compense entièrement la baisse d'une unité dans une autre dimension. D'autre part, si $\gamma > 1$, la compensation n'est pas totale. Mais si $\gamma \rightarrow \infty$, les biens sont complémentaires (substituabilité nulle). La substituabilité des attributs admet l'hypothèse selon laquelle la détérioration d'un attribut peut être compensée par l'amélioration d'un autre attribut. Cette hypothèse est fortement rejetée parce qu'elle contredit une hypothèse forte de la théorie de mesurage reposant sur l'incommensurabilité des dimensions.

Le principal avantage de cet indice vient du fait qu'elle conduit à un préordre complet. D'autre part, cet indice permet de résumer l'information globale et d'éviter la confusion qui peut être liée à la présence d'une pléthore d'indicateurs individuelles. Un indice synthétique fournit également une meilleure communication car elle permet de se focaliser sur le problème en évitant la pluralité d'indices.

Le principal inconvénient d'un indice composite est lié au fait qu'une grande quantité d'informations est perdue pendant le processus d'agrégation. De ce fait, l'image que ce dernier fournit peut parfois être trompeuse et ainsi aboutir à des politiques actions simplistes et erronées. D'autre part, un indice composite ne retranscrit que peut d'informations comparé aux informations obtenues sur les indicateurs pris séparément sur chaque variable du bien-être.

2.3.3.7 Les approches basées sur les données individuelles

Les mesures alternatives sont généralement basées sur les données individuelles. Nous avons notamment l'approche de la fonction de distance [voir Deutsch et Silber, 2005; Silber, 2007], l'approche de la théorie de l'information ou approche d'entropie [voir Theil, 1967; Foster *et al.*, 1984; Maasoumi, 1986, 1993; Cowell, 1977; Cowell et Kuga, 1981a,b; Lugo et Maasoumi, 2008] et l'approche d'inertie [see Klasen, 2000; Sahn et Stifel, 2000, 2003].

a) L'approche par la fonction distance

La fonction distance est un concept largement utilisé dans l'analyse d'efficacité [voir Coelli *et al.*, 1998]. Lovell *et al.* [1990] furent les premiers à appliquer ce concept dans le cadre de l'analyse du comportement des ménages. Deutsch et Silber [2005] l'a ensuite utilisé dans le cadre du mesurage multidimensionnel de la pauvreté. Cette approche consiste à mesurer de quel montant ajuster l'ensemble des attributs d'un individu pour atteindre un niveau de bien-être de référence. Cela se fait en construisant une mesure de la distance entre un vecteur de biens (ensembles de ressources) d'un individu et un *vecteur de référence*.

Formellement, on définit $L(y)$ comme étant l'ensemble d'entrée de tous les vecteurs d'entrée x qui peuvent produire le vecteur de sortie y .

$$L(y) = \{x \text{ t.q. } x \text{ peut produire } y\}. \quad (2.26)$$

La fonction distance d'entrée $D_{in}(x)$ est définie par l'équation :

$$D_{in}(x, y) = \max \left\{ \tau \text{ t.q. } \left(\frac{x}{\tau} \right) \in L(y) \right\} \quad (2.27)$$

Coelli *et al.* [1998] ont démontré la fonction distance d'entrée respecte les quatre propriétés suivantes :

- (i) $D_{in}(x, y)$ est croissante en x et décroissante en y .
- (ii) $D_{in}(x, y)$ est linéairement homogène en x .
- (iii) Si $x \in L(y)$ alors $D_{in}(x, y) \geq 1$.
- (iv) $D_{in}(x, y) = 1$ si x appartient à la "frontier" de $L(y)$ (l'isoquante de y).

Deutsch et Silber [2005] ont proposé dans leur article une estimation de l'indice du niveau de vie standard sur la base des données sur la possession de biens durables en s'appuyant sur la théorie de la fonction distance. Ils définissent le vecteur des ressources $x = (r_1, r_2, \dots, r_N) \in \mathbb{R}_+^N$ et le vecteur des fonctionnements $u = (u_1, u_2, \dots, u_M) \in \mathbb{R}_+^M$. Alors chaque individu est représenté par le couple (r^i, u^i) , $i \in I$. Un indice théorique du niveau de vie standard (standard of living index) SL est ensuite estimé en utilisant un indice de quantité d'entrée de Malmquist [Coelli *et al.*, 1998] selon l'équation (2.28) :

$$SL(u, r^s, r^t) = \frac{D_{in}(u, r^s)}{D_{in}(u, r^t)} \quad (2.28)$$

où r^s et r^t sont deux vecteurs de ressources différents et D_{in} est une fonction distance d'entrée définie selon l'équation (2.27). L'indice de Malmquist permet de fournir un *ensemble référence* permettant de juger de l'ampleur relative des deux vecteurs de ressources. Cet ensemble de référence est le isoquante $L(u)$ et le plus radicalement éloigné de r^i est de $L(u)$ supérieur à son niveau de vie, pour que r^i soit réduit progressivement pour revenir sur l'ensemble de référence $L(u)$. Il y a cependant une difficulté parce que l'indice de Malmquist dépend généralement de u . Pour contourner la difficulté, Deutsch et Silber [2005] ont proposé de se débarrasser de u en traitant tous les individus de manière égale et suppose que chaque individu a le même niveau de fonctionnements : une unité pour chaque fonctionnement. Soit e un tel vecteur de fonctionnements (un M -vecteur de un). Ainsi, l'ensemble référence devient $L(e)$ et borne les vecteurs ressources ci-dessous. Les individus muni de vecteurs ressources sur $L(e)$ se partagent le plus bas niveau de vie avec une valeur d'index de l'unité, tandis que les individus muni de vecteurs ressources larges auront alors des niveaux de vie avec des valeurs de l'indice supérieur à l'unité.

Pour estimer la fonction distance, Deutsch et Silber [2005] définissent $\lambda = 1/r_N$ et un $(N - 1)$ -vecteur $z = \{z_j\} = \{r_j/r_N\}$ avec $j = 1, \dots, (N - 1)$. De là :

$$D_{in}(z, e) = (1/x_N) \cdot D_{in}(x, e) \quad (2.29)$$

Comme $D_{in}(x, e) \geq 1$, on obtient :

$$(1/r_N) \leq D_{in}(x, e) \quad (2.30)$$

Cela implique l'écriture suivante :

$$(1/r_N) = D_{in}(x, e) \cdot \exp(\varepsilon), \quad \varepsilon \leq 0. \quad (2.31)$$

En supposant que $D_{in}(z, e)$ a une *forme fonctionnelle translog*, on obtient :

$$\ln(1/r_N) = \alpha_{00} + \sum_{j=1}^{N-1} \alpha_{j0} \ln(z_j) + \frac{1}{2} \sum_{j=1}^{N-1} \sum_{k=1}^{N-1} \alpha_{jk} \ln(z_j) \ln(z_k) + \varepsilon. \quad (2.32)$$

L'estimation des coefficients α_{00} , α_{j0} et α_{jk} (avec $j = 1, \dots, (N - 1)$) peut être obtenue en utilisant la méthode des moindres carrés ou la méthode du maximum de vraisemblance tandis que la fonction distance d'entrée $D_{in}(z^i, e)$ pour chaque individu x_i est donnée par la transformation :

$$D_{in}(z^i, e) = \exp\{\max(\varepsilon_I) - \varepsilon_i\} \quad (2.33)$$

Le niveau de vie d'un individu x_i sera alors obtenu en divisant $D_{in}(z^i, e)$ par la valeur de la distance minimale observée— laquelle est égale à 1.

b) L'approche de la théorie de l'information

La théorie de l'information a été originellement développée par Claude Shannon en tant que discipline au sein de la théorie mathématique de la communication. L'objectif était de déterminer combien de données peuvent être transmises via un canal sans pertes ou erreurs significatives [voir [Shannon, 1948](#)]. Elle a ensuite été utilisée pour aborder la question de l'agrégation d'attributs pour les évaluations du bien-être dans le cadre de la mesure unidimensionnelle et multidimensionnelle de l'inégalité [cf. [Theil, 1967](#); [Maasoumi, 1986](#)]. [Maasoumi \[1993\]](#) en expose le principe de base et définit l'entropie comme une mesure de l'incertitude ou de la volatilité associée à une variable aléatoire. Il s'agit de définir des classes de mesures pour capter les divergences entre distributions. Cowell et Kuga [[Cowell, 1977](#); [Cowell et Kuga, 1981a,b](#)] étendent la famille d'indices d'inégalité basée sur la théorie de l'information en des mesures d'entropie généralisées (*Generalized Entropy measures*) :

$$GE_\gamma = \frac{1}{\gamma(1-\gamma)} \frac{1}{n} \sum_{i=1}^n \left[1 - \left(\frac{r_i}{\bar{r}} \right)^\gamma \right] \quad (2.34)$$

où r_i est le revenu du ménage x_i et $\bar{r} = \frac{1}{n} \sum_{i=1}^n r_i$ est le revenu moyen. $\gamma \in] - \infty; +\infty[$ est un paramètre de capture de la sensibilité d'un indice particulier GE , aux différentes parties de la distribution. Plus γ est petit, plus la mesure de la sensibilité de la borne inférieure est élevée, ce sont les pauvres. Le premier indice de Theil T_1 et le deuxième indice de Theil T_2 sont obtenus comme suit :

- Pour $\gamma = 1$, $GE_1 = T_1 = \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{r}_i}{\bar{\mathbf{r}}} \ln \left(\frac{\mathbf{r}_i}{\bar{\mathbf{r}}} \right)$
- Pour $\gamma = 0$, $GE_0 = T_2 = \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{\bar{\mathbf{r}}}{\mathbf{r}_i} \right)$

Dans le cadre de la mesure multidimensionnelle de l'inégalité, [Maasoumi \[1986\]](#) utilise la théorie de l'information tant dans l'agrégation des attributs (pour obtenir un indice du bien-être de chaque individu) que dans l'agrégation des individus pour obtenir la mesure de l'inégalité. La fonction distance $\mathcal{D}_\gamma(\cdot)$ est définie comme étant la moyenne pondérée des divergences d'entropie relative entre $\mathbf{f}_i = \{\mathbf{f}_1(\mathbf{z}, X), \mathbf{f}_2(\mathbf{z}, X), \dots, \mathbf{f}_n(\mathbf{z}, X)\}$ et $f = \{f(z^1, x_i^1), f(z^2, x_i^2), \dots, f(z^m, x_i^m)\}$ formulée comme suite :

$$\mathcal{D}_\gamma(\mathbf{f}_i \parallel X; \lambda) = \sum_{j=1}^m \lambda_j \frac{1}{\gamma(1-\gamma)} \sum_{i=1}^n \mathbf{f}_i(\mathbf{z}, X) \left[1 - \left(\frac{\mathbf{f}_i(\mathbf{z}, X)}{f(z^j, x_i^j)} \right)^\gamma \right] \quad (2.35)$$

où λ_j est le poids associé à la distance d'entropie généralisée de chaque attribut. La minimisation de $\mathcal{D}_\gamma(\mathbf{f}_i \parallel X; \lambda)$ sous la contrainte $\sum_{i=1}^n \mathbf{f}_i(\mathbf{z}, X)$ fournit les *fonctions d'agrégation optimales* suivantes :

$$\mathbf{f}_i(\mathbf{z}, X) = \left(\sum_{j=1}^m \lambda_j (f(z^j, x_i^j))^\gamma \right)^{1/\gamma} \text{ lorsque } \gamma \neq 0 \quad (2.36)$$

$$\mathbf{f}_i(\mathbf{z}, X) = \prod_{j=1}^m (f(z^j, x_i^j))^{\lambda_j} \text{ lorsque } \gamma = 0 \quad (2.37)$$

où γ est un paramètre qui détermine le niveau de substitution entre les attributs dans la fonction d'agrégation. [Maasoumi](#) propose ensuite d'utiliser une mesure d'entropie généralisée (Generalized Entropy measure) sur les indices du bien-être résultants :

$$GEM_\gamma = \frac{1}{\gamma(1-\gamma)} \frac{1}{n} \sum_{i=1}^n \left[1 - \left(\frac{\mathbf{f}_i(\mathbf{z}, X)}{\bar{\mathbf{f}}_i(\mathbf{z}, X)} \right)^\gamma \right] \quad (2.38)$$

Cette mesure est accusée de ne pas satisfaire la version multidimensionnelle du principe de transfert de Pigou-Dalton [[Lugo et Maasoumi, 2008](#)]. [Lugo et Maasoumi \[2008\]](#) contournent le problème en utilisant une approche similaire à celle de [Maasoumi \[1986\]](#), en se basant sur l'approche de la théorie d'information, pour dériver des indices multidimensionnels de la pauvreté. Deux approches ont ainsi été érigées par [Lugo et Maasoumi \[2008\]](#) : le *seuil de pauvreté agrégé* (aggregate poverty line) et le *seuil de pauvreté composé* (component poverty line).

Le seuil de pauvreté est le *seuil de pauvreté agrégé* \mathbf{z} associé aux fonctions d'agrégation de la théorie de l'information. Ce seuil est donné par les équations (2.36) et

(2.37) :

$$\mathbf{z} = \left(\sum_{j=1}^m \lambda_j (z^j)^\gamma \right)^{1/\gamma} \text{ lorsque } \gamma \neq 0 \quad (2.39)$$

$$\mathbf{z} = \prod_{j=1}^m (z^j)^{\lambda_j} \text{ lorsque } \gamma = 0 \quad (2.40)$$

Le seuil de pauvreté du j^{e} attribut z^j joue un rôle important dans la définition de $\mathbf{f}_i(\mathbf{z}, X)$, laquelle incorpore les mêmes poids et relations entre les attributs considérés individuellement. La fonction d'agrégation F de l'équation (2.5) est définie par :

$$\mathbf{P}(\mathbf{z}, X) = \frac{1}{n} \sum_{i=1}^n \left[\max \left\{ \frac{\mathbf{z} - \mathbf{f}_i(\mathbf{z}, X)}{\mathbf{z}}; 0 \right\} \right]^\gamma \mathbf{I}(\mathbf{z} > \mathbf{f}_i(\mathbf{z}, X)) \quad (2.41)$$

Dans l'approche du seuil de pauvreté composé, on cherche à obtenir les fonctions de synthèse de

$$\delta_{\mathbf{x}_i}^j = \frac{z^j - f(z^j, x_i^j)}{z^j} \quad (2.42)$$

au lieu de $f(z^j, x_i^j)$. Les fonctions de synthèse $\delta_{\mathbf{x}_i}^j$ peuvent être interprétées comme les insuffisances au seuil (*shortfalls to threshold*) où $0 \leq \delta_{\mathbf{x}_i}^j \leq 1$ pour les *pauvres* et $\delta_{\mathbf{x}_i}^j \leq 0$ pour les *non-pauvres* [Bourguignon et Chakravarty, 2003]. L'indicateur d'insuffisance est alors donné par les équations (2.43) et (2.44) :

$$g(\delta_{\mathbf{x}_i}) = \left(\sum_{j=1}^m \lambda_j \left(\frac{z^j - f(z^j, x_i^j)}{z^j} \right)^\gamma \right)^{1/\gamma} \text{ lorsque } \gamma \neq 0 \quad (2.43)$$

$$g(\delta_{\mathbf{x}_i}) = \prod_{j=1}^m \left(\frac{z^j - f(z^j, x_i^j)}{z^j} \right)^{\lambda_j} \text{ lorsque } \gamma = 0 \quad (2.44)$$

Les λ_j sont des poids positifs associés à chaque j^{e} me insuffisance (*shortfall*), γ est le niveau de substitution entre les insuffisances ; plus γ est élevé, plus le degré de substitution entre les insuffisances est petit. Lorsque $\gamma \rightarrow \infty$, les déprivations relatives sont complémentaires (substituabilité nulle) ; lorsque $\gamma > 1$, la compensation n'est pas totale et lorsque $\gamma = 1$, la compensation est totale (substituabilité infinie). La fonction de pauvreté individuelle est ensuite donnée par l'équation (2.45) et la mesure multidimensionnelle par l'équation (2.46).

$$\mathbf{g}_i(\mathbf{z}, X) = \left[\sum_{j=1}^m \lambda_j \max \{ \delta_{\mathbf{x}_i}^j; 0 \}^\gamma \right]^{1/\gamma} \quad (2.45)$$

$$\mathbf{P}(\mathbf{z}, X) = \frac{1}{n} \sum_{i=1}^n \mathbf{g}_i^\alpha(\mathbf{z}, X) \mathbf{I}(\mathbf{g}_i(\mathbf{z}, X) > 0) \quad (2.46)$$

qui est le α^e moment de la distribution de $g(\delta) = \{g(\delta_{x_1}), g(\delta_{x_2}), \dots, g(\delta_{x_n})\}$. La mesure multidimensionnelle de la pauvreté à partir de la composante ligne de pauvreté (*component poverty line*) est donc :

$$\mathbf{P}_{dur}(\mathbf{z}, X) = \frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^m \lambda_j \max \left\{ \frac{z^j - f(z^j, x_i^j)}{z^j}; 0 \right\}^\gamma \right]^{\alpha/\gamma} \mathbf{I}(\mathbf{g}_i(\mathbf{z}, X) > 0) \quad (2.47)$$

et

$$\mathbf{P}_{dur}(\mathbf{z}, X) = \frac{1}{n} \sum_{i=1}^n \left[\prod_{j=1}^m \max \left\{ \frac{z^j - f(z^j, x_i^j)}{z^j}; 0 \right\}^{\lambda_j} \right]^\alpha \mathbf{I}(\mathbf{g}_i(\mathbf{z}, X) > 0) \quad (2.48)$$

Le lecteur intéressé pourra consulter les ouvrages de [Shannon \[1948\]](#); [Theil \[1967\]](#); [Foster et al. \[1984\]](#); [Maasoumi \[1986, 1993\]](#); [Lugo et Maasoumi \[2008\]](#) pour plus de détails. Toutefois, la faiblesse de cette démarche est liée au problème d'indétermination des paramètres adéquats aux mesures proposées. À ce problème, s'ajoute celui d'une détermination "moins arbitraire" du système de pondération $\{\lambda_j\}_j$ des attributs. Les méthodes d'inertie apportent des solutions à cette seconde préoccupation.

c) L'approche d'inertie

L'approche d'inertie [see [Asselin, 2009](#)] est une approche paramétrique de l'indicateur composite de pauvreté qui est principalement basée sur des techniques d'analyse multivariées. Une brève revue de la littérature sur les approches d'inertie pour la mesure de la pauvreté nous a permis de mettre en lumière les techniques suivantes.

L'analyse en composant principal (ACP)

[Filmer et Pritchett \[2001\]](#) a construit un indice linéaire de la richesse basé sur des actifs des variables indicatrices de propriété en utilisant des données provenant de l'Inde. Ils ont utilisé l'ACP pour obtenir la première composante principale normalisée de la matrice de variance-covariance des actifs des ménages observés (également appelé *poids*) pour les indicateurs d'actifs. Ils ont conclu qu'en l'absence de données sur les dépenses de consommation, l'application de l'ACP pour calculer les indices d'actifs est une alternative cohérente et stable [voir [Klasen, 2000](#); [Filmer et Pritchett, 1997](#); [Asselin, 2009](#); [Kabubo-Mariara et al., 2011](#), pour plus de détails et des applications].

Analyse factorielle (AF)

Sahn et Stifel [2003] a utilisé l'analyse factorielle pour évaluer le potentiel d'un indice reposant sur l'actif comme un indicateur du bien-être économique des ménages. Dans le cas d'une approche par les capacités, Silber [2007] a fait valoir que l'analyse des facteurs fournirait un cadre théorique pour expliquer les fonctionnements (observés) au moyen de capacités représentées par les facteurs latents, mais un tel modèle ne peut pas être utilisé pour expliquer les variables latentes [voir Hirschberg *et al.*, 1991; Nolan et Whelan, 1996; Asselin, 2009, pour plus de détails].

L'analyse en correspondance multiple (ACM)

Booyesen *et al.* [2007] a utilisé ACM pour construire des indicateurs composites de la pauvreté en se basant sur les actifs des ménages. ACM est préférable lorsque nous sommes en présence de données catégoriques et il peut facilement combiner des variables quantitatives et les variables catégorielles. ACM est aussi utile pour la visualisation des données et des variables sur le même graphique [voir Asselin, 2009, pour plus de détails].

Cluster analysis (CA)

Ferro-Luzzi *et al.* [2006] a utilisé CA pour agréger les individus selon leurs similarités au regard des différents scores de privations multiples. Hirschberg *et al.* [1991] a proposé des méthodes statistiques de CA pour explorer différentes façons et niveaux de regroupement des 23 attributs divers tels que les droits politiques, les libertés civiles, de l'espérance de vie, l'alphabétisation, le produit intérieur réel, etc. Ensuite, ils a utilisé une analyse typologique pour visualiser comment les unités de bien-être (pays) se regroupent sur la base des attributs considérés, d'une manière (métrique) conforme à celle utilisée dans le but de regrouper les attributs.

2.3.3.8 Les approches par l'intersection et par l'union

L'approche basée sur l'intersection [voir Duclos *et al.*, 2006, pour plus de détails] considère qu'un ménage \mathbf{x}_i est pauvre s'il n'arrive à satisfaire aucun des besoins essentiels. Le ménage \mathbf{x}_i est *pauvre* s'il l'est sur tous les attributs X_j ($j \in \{1, \dots, m\}$). Selon ce point de vue, la fonction $f_i(\mathbf{z}, X)$ est donnée par l'équation suivante :

$$f_i(\mathbf{z}, X) = \begin{cases} \tau > 0, & \text{si } x_i^j < z_j, \forall j \in \{1, \dots, m\}; \\ 0, & \text{sinon.} \end{cases} \quad (2.49)$$

L'approche basée sur l'union considère qu'un ménage \mathbf{x}_i est pauvre s'il n'arrive pas à satisfaire *au moins un* des besoins essentiels. Le ménage \mathbf{x}_i est *pauvre* s'il l'est pour *au moins un* des attributs X_j ($j \in \{1, \dots, m\}$). Selon ce point de vue, la

fonction $f_i(\mathbf{z}, X)$ est donnée par l'équation suivante :

$$f_i(\mathbf{z}, X) = \begin{cases} 0, & \text{si } x_i^j \geq z_j, \forall j \in \{1, \dots, m\}; \\ \tau > 0, & \text{sinon.} \end{cases} \quad (2.50)$$

Remarque 2.3.2 *Les approches de la pauvreté par l'intersection et par l'union traitent de la même manière les individus qui ne parviennent pas à satisfaire un seul besoin essentiel ou plusieurs. Elles ne permettent pas alors de tenir compte de la profondeur et de la sévérité de la pauvreté. Bourguignon et Chakravarty [2003] conseillent de ne pas sommer le nombre de pauvres dans les différents D_j (avec $j \in \{1, \dots, m\}$) pour avoir le nombre total de pauvres car cela entraînerait un "double comptage". Ils proposent alors de calculer l'incidence de la pauvreté à partir l'approche par l'union, c'est-à-dire, du nombre d'individus ayant au moins une insuffisance dans un attribut.*

2.3.3.9 L'approche par la dominance stochastique bidimensionnelle

Les approches par la dominance stochastique bidimensionnelle sont des extensions multidimensionnelles de l'approche de la dominance stochastique développée par Atkinson [1987]; Foster et Shorrocks [1988a,b]; Zheng [1999, 2001] dans le cadre unidimensionnel. Elles permettent de procéder à des comparaisons ordinales de pauvreté. L'approche de dominance stochastique multidimensionnelle cherche à définir les conditions nécessaires pour que ces comparaisons restent valables à l'intérieur d'un intervalle plausible de variation de \mathbf{z} et pour une famille de mesures de pauvreté. Bien que les applications de la dominance stochastique dans le cadre unidimensionnel sont variées [voir Atkinson, 1987; Foster et Shorrocks, 1988a,b; Zheng, 1999, 2001], les applications dans un contexte multidimensionnel se limitent généralement dans un contexte bidimensionnel [voir Atkinson et Bourguignon, 1982; Bourguignon et Chakravarty, 2002; Yélé, 2007, 2008; Sami et Lahga, 2008] et demeurent donc une piste de recherche importante.

Bourguignon et Chakravarty [2002] considèrent un indice de pauvreté satisfaisant les axiomes (voir sous-section 2.3.4) de *concentration*, de *symétrie*, de *décomposabilité par sous-groupe*, du *principe de la population* et deux fois *différentiable* (ce qui remplace l'axiome de continuité) sous l'hypothèse que la limite supérieure des seuils de pauvreté pour chaque attribut est fixée. Ils considèrent un continuum d'individus ainsi qu'une distribution de deux attributs $\mathbf{x}_i = (x_i^1, x_i^2) = (a_1, a_2)$. Soient deux fonctions de distributions cumulatives ψ et ψ^* définies sur $[0, z_1^*] \times [0, z_2^*]$, les mesures de pauvreté associées à ces distributions sont données par les équations (2.51) et (2.52) suivantes :

$$\mathbf{P}(\mathbf{z}, \psi) = \int_0^{z_1^*} \int_0^{z_2^*} f_i(z^1, z^2; a_1, a_2) d\psi(a_1, a_2) \quad (2.51)$$

$$\mathbf{P}(\mathbf{z}, \psi^*) = \int_0^{z_1^*} \int_0^{z_2^*} \mathbf{f}_i(z^1, z^2; a_1, a_2) d\psi^*(a_1, a_2) \quad (2.52)$$

où $\mathbf{f}_i(z^1, z^2; a_1, a_2)$ est la fonction de pauvreté individuelle associée à un individu \mathbf{x}_i ayant (a_1, a_2) comme attributs. L'objectif est de comparer les deux mesures de pauvreté obtenues par les équations (2.51) et (2.52). Pour cela, on s'intéresse à la différence entre les deux mesures :

$$\Delta\mathbf{P}(\mathbf{z}) = \mathbf{P}(\mathbf{z}, \psi) - \mathbf{P}(\mathbf{z}, \psi^*) = \int_0^{z_1^*} \int_0^{z_2^*} \mathbf{f}_i(z^1, z^2; a_1, a_2) d\Phi(a_1, a_2) \quad (2.53)$$

avec $\Phi(a_1, a_2) = \psi(a_1, a_2) - \psi^*(a_1, a_2)$. On dira que la distribution ψ **domine** ψ^* si $\Delta\mathbf{P}(\mathbf{z}) < 0$.

Bourguignon et Chakravarty [2002] ont défini trois classes de mesures de pauvreté avec deux attributs substituables, complémentaires, ou indépendants.

★ La substituabilité

La dominance stochastique de premier ordre dans chaque attribut de la pauvreté peut être appliquée lorsque les deux attributs sont *substituables*. Dans ce cas, on a :

$$\frac{\partial^2 \mathbf{f}_i(z^1, z^2; a_1, a_2)}{\partial a_1 \partial a_2} > 0 \text{ et } \Delta\mathbf{P}(a_j) = \int_0^{a_j} d\Phi_j(b_j) \leq 0, \forall a_j, b_j \leq z_j^* \quad (2.54)$$

Bourguignon et Chakravarty [2002] ont démontré que la substituabilité des attributs reflète l'intersection des multiples dimensions de la pauvreté. Ainsi, la dominance du premier ordre à travers l'intersection des deux dimensions de la pauvreté est donnée par l'équation (2.55) suivante :

$$\Delta\mathbf{P}(\mathbf{x}_i) = \int_0^{a_1} \int_0^{a_2} d\Phi(b_1, b_2) \leq 0, \forall a_1 \leq z_1^* \text{ et } a_2 \leq z_2^* \quad (2.55)$$

★ La complémentarité

La condition de robustesse de premier ordre sur chaque attribut de la pauvreté peut être appliquée lorsque les deux attributs sont *complémentaires*. Dans ce cas, on a :

$$\frac{\partial^2 \mathbf{f}_i(z^1, z^2; a_1, a_2)}{\partial a_1 \partial a_2} < 0 \text{ et } \Delta\mathbf{P}(a_j) = \int_0^{a_j} d\Phi_j(b_j) \leq 0, \forall a_j, b_j \leq z_j^* \quad (2.56)$$

Ainsi, la dominance stochastique du premier ordre, avec la condition de robustesse, à travers l'union des deux dimensions de la pauvreté est donnée par l'équation (2.57) suivante :

$$\Delta\mathbf{P}(\mathbf{x}_i) = \sum_{j=1;2} \int_0^{a_j} d\Phi_j(b_j) - \int_0^{a_1} \int_0^{a_2} d\Phi(b_1, b_2) \leq 0, \forall a_1 \leq z_1^* \text{ et } a_2 \leq z_2^* \quad (2.57)$$

★ L'indépendance

Lorsque les deux attributs sont *indépendants*, la dominance stochastique vérifie les conditions données par l'équation (2.58) suivante :

$$\frac{\partial^2 \mathbf{f}_i(z^1, z^2; a_1, a_2)}{\partial a_1 \partial a_2} = 0 \text{ et } \Delta \mathbf{P}(a_j) = \int_0^{a_j} d\Phi_j(b_j) \leq 0, \forall a_j, b_j \leq z_j^* \quad (2.58)$$

En 2006, les conditions de robustesse ont été établies par [Duclos *et al.* \[2006\]](#) dans le cadre de leurs travaux sur les comparaisons de la pauvreté utilisant les indicateurs multidimensionnels du bien-être. Ils parviennent ainsi à démontrer, en particulier, comment vérifier si les comparaisons de la pauvreté sont *robustes* par rapport aux procédures d'agrégation et au choix des lignes de pauvreté multidimensionnelle (c'est-à-dire, le vecteur ξ). Ainsi, [Duclos *et al.* \[2006\]](#) commencent par définir un indicateur de bien-être individuel donné par l'équation (2.59) :

$$\varphi(a_1, a_2) : \mathbb{R}^2 \longrightarrow \mathbb{R} \left| \frac{\partial \varphi(a_1, a_2)}{\partial a_1} \geq 0, \frac{\partial \varphi(a_1, a_2)}{\partial a_2} \geq 0 \quad (2.59)$$

[Duclos *et al.* \[2006\]](#) supposent ensuite qu'il existe une frontière *inconnue* (qu'il faut déterminer) séparant les pauvres de celle non pauvres. Cette frontière est implicitement définie par $\varphi(a_1, a_2) = 0$. Partant de cette hypothèse, ils définissent l'ensemble des individus pauvres par l'équation (2.60) :

$$\Lambda(\varphi(a_1, a_2)) = \{(a_1, a_2) | \varphi(a_1, a_2) \leq 0\} \quad (2.60)$$

Une mesure de pauvreté bidimensionnelle respectant l'axiome de la *décomposabilité par sous-groupe* peut donc être donnée par :

$$\mathbf{P}(\varphi) = \int \int_{\Lambda(\varphi)} \mathbf{f}_i(\varphi; a_1, a_2) d\Phi(a_1, a_2) \quad (2.61)$$

où $\mathbf{f}_i(\varphi; a_1, a_2)$ est la contribution d'un individu $\mathbf{x}_i = (a_1, a_2)$ à la pauvreté globale. A partir de l'axiome de focalisation, cette fonction respecte les conditions définies par l'équation (2.62) :

$$\mathbf{f}_i(\varphi; a_1, a_2) \begin{cases} \geq 0 & \text{si } \varphi(a_1, a_2) \leq 0; \\ = 0 & \text{sinon.} \end{cases} \quad (2.62)$$

L'incidence bidimensionnelle de la pauvreté est obtenue lorsque $\mathbf{f}_i(\varphi; a_1, a_2) = 1$ si $\varphi(a_1, a_2) \leq 0$.

En considérant $\psi(a_1)$ comme une fonction de distribution de l'attribut a_1 , la courbe de dominance stochastique unidimensionnelle est donnée par l'équation (2.63) :

$$\mathbf{P}^{\alpha_1}(z^1; a_1) = \int_0^{z^1} (z^1 - a_1)^{\alpha_1} d\psi(a_1) \quad \text{pour un entier } \alpha_1 \geq 0. \quad (2.63)$$

L'équation (2.63) représente une moyenne pondérée des écarts de pauvreté $(z^1 - a_1)$ de paramètre α_1 , et représente aussi l'indice de pauvreté populaire de Foster *et al.* [1984].

Pour les besoins de l'analyse de robustesse, Duclos *et al.* [2006] considèrent une surface de dominance stochastique bidimensionnelle définie par l'équation (2.64) qui est l'extension multidimensionnelle de la classe de mesures FGT :

$$\mathbf{P}^{\alpha_1, \alpha_2}(z^1, z^2; a_1, a_2) = \int_0^{z^1} \int_0^{z^2} (z^1 - a_1)^{\alpha_1} (z^2 - a_2)^{\alpha_2} d\psi(a_1, a_2). \quad (2.64)$$

Pour les entiers $\alpha_1 \geq 0$ et $\alpha_2 \geq 0$. Une seconde caractéristique importante de la surface de dominance est qu'elle est influencée par la covariance entre a_1 et a_2 (qui sont les mesures du bien-être *uniattributs*) parce que l'intégrale est *multiplicative*. Ce qui permet alors de réécrire l'équation (2.64) comme suit :

$$\begin{aligned} \mathbf{P}^{\alpha_1, \alpha_2}(z^1, z^2; a_1, a_2) &= \int_0^{z^1} (z^1 - a_1)^{\alpha_1} d\psi(a_1) \int_0^{z^2} (z^2 - a_2)^{\alpha_2} d\psi(a_2) \\ &\quad + \text{cov}[(z^1 - a_1)^{\alpha_1}, (z^2 - a_2)^{\alpha_2}] \end{aligned} \quad (2.65)$$

$\mathbf{P}^{0,0}(z^1, z^2; a_1, a_2)$ représente l'incidence bidimensionnelle de pauvreté, c'est-à-dire le pourcentage de la population pauvre selon les deux attributs simultanément. $\mathbf{P}^{1,0}(z^1, z^2; a_1, a_2)$ est la mesure agrégée du déficit de pauvreté de a_1 des individus pauvres au regard du deuxième attribut. Tandis que $\mathbf{P}^{1,1}(z^1, z^2; a_1, a_2)$ agrège le produit des déficits de pauvreté, normalisé par la taille de la population.

2.3.4 Caractérisation axiomatique

Les caractérisations axiomatiques ont été préconisées afin de permettre une conceptualisation multidimensionnelle de la pauvreté et d'agréger des dimensions telles en un indice composite (comme dans le cas des indices monétaires) en conformité avec les propriétés (ou axiomes) que l'indice de la pauvreté doivent satisfaire. Les premiers auteurs à adopter une approche axiomatique dans le contexte d'indices multidimensionnels de pauvreté fût Chakravarty *et al.* [1998], suivi de Tsui [2002] après ces précédents travaux portant sur l'axiomatisation des indices multidimensionnels d'inégalité Tsui [1995], et les travaux de Bourguignon et Chakravarty [2002, 2003].

2.3.4.1 Quelques axiomes et leurs interprétations

Une mesure multidimensionnelle de la pauvreté $\mathbf{P}(z, X)$ doit respecter un certain nombre d'axiomes qui, pour la plupart, sont une extension multidimensionnelle d'un ensemble de propriétés désirables pour les indices de pauvreté. Ces

axiomes ont un impact direct sur les fonctions d'agrégation qui permettent de *mesurer* la magnitude de la pauvreté multidimensionnelle. Les fonctions $G(\cdot)$ et $F(\cdot)$ dépendront donc explicitement des axiomes que la mesure de pauvreté devra absolument respecter. La littérature [Zheng, 1997; Chakravarty *et al.*, 1998; Tsui, 2002; Bourguignon et Chakravarty, 2003; Sami, 2005] renferme une panoplie d'axiomes relatifs aux mesures multidimensionnelles de pauvreté. Tandis que certains de ces axiomes sont nouveaux, d'autres par contre sont une simple généralisation de ceux inhérents à la construction d'une mesure unidimensionnelle de pauvreté.

Axiome 2.3.1 La continuité (CN) : Cette axiome stipule que la mesure de pauvreté ne doit pas être très sensible à une variation marginale de la quantité d'un attribut. Il évite que des variations mineures de mesure (parfois dues à des erreurs d'observation) sur les quantités des attributs n'engendrent des changements brutaux des comparaisons de la pauvreté. Formellement, pour tout $\mathbf{z} \in Z$, $\mathbf{P}(\mathbf{z}, X)$ est continue sur X .

Axiome 2.3.2 Symétrie (SM) : L'axiome de symétrie (ou l'anonymat) stipule qu'aucune caractéristique, autre que les quantités x_i^j d'attribut $j \in \{1, \dots, m\}$ utilisées pour définir l'indice multidimensionnel de pauvreté, ne doit en aucun cas influencer les résultats d'analyse. Autrement dit, tout indicateur non pertinent (tels que son nom de famille, sa taille, l'identité des individus) ne doit pas être pris en compte et une permutation des paniers de biens entre individus ne change pas la mesure de la pauvreté. Formellement, pour tout X et pour tout $\mathbf{z} \in Z$, $\mathbf{P}(\mathbf{z}, X) = \mathbf{P}(\mathbf{z}, \Pi_\sigma X)$, où Π_σ représente une matrice de permutation (voir définition 2.3.2).

Axiome 2.3.3 La focalisation (FO) : La focalisation ou la concentration⁶ stipule que la mesure de pauvreté $\mathbf{P}(\mathbf{z}, X)$ reste inchangée si la situation d'un individu \mathbf{x}_i s'améliore sur un attribut X_j sur lequel il n'est pas pauvre (attribut *non-maigre*). Formellement, soit $X = [x_i^j]$ and $\tilde{X} = [y_i^j]$, si pour tout $\mathbf{x}_i \in \Omega$ tel que $y_i^j = x_i^j + \varepsilon$ et $x_i^j \geq z^j$, avec $\varepsilon > 0$ alors $\mathbf{P}(\mathbf{z}, \tilde{X}) = \mathbf{P}(\mathbf{z}, X)$.

Bourguignon et Chakravarty [2003] a proposé deux versions de l'axiome de focalisation : *la focalisation forte* et *la focalisation faible*.

Axiome 2.3.4 La focalisation forte (SF) : L'axiome de focalisation forte implique la non compensation (ou non substituabilité) entre les attributs "maigres" et "non maigres" des individus pauvres. En d'autres termes, des mesures de la pauvreté ne devraient pas tenir compte de l'amélioration du bien-être des individus qui ne sont pas pauvres. Formellement, soit $X = [x_i^j]$ et $\tilde{X} = [y_i^j]$, (a) si pour tout $\mathbf{x}_i \in \Omega$ tel que $x_i^j \geq z^j$, $y_i^j = x_i^j + \varepsilon$ (avec $\varepsilon > 0$), (b) $x_k^j = y_k^j$ pour tout $k \neq i$, et (c) $x_i^l = y_i^l$ pour tout $l \neq j$ et pour tout \mathbf{x}_i , $i, k \in \{1, \dots, n\}$ alors $\mathbf{P}(\mathbf{z}, \tilde{X}) = \mathbf{P}(\mathbf{z}, X)$.

6. La terminologie "concentration" est utilisée par Sami et Lahga [2008].

Pour illustrer [voir Chakravarty *et al.*, 1998] cet axiome, considérons les trois dimensions X_1 (capabilités en termes de santé), X_2 (capabilités en termes d'éducation) et X_3 (capabilités en termes de conditions de vie) utilisés pour définir l'indice de pauvreté humaine (IPH). Supposons que le vecteur des seuils de pauvreté soit donné par $\mathbf{z} = (0.8; 0.5; 0.6)$. Soit $\mathbf{x}_i = (0.7; 0.95; 0.6)$ le vecteur des attributs d'un individu i . Bien que son niveau d'éducation soit largement supérieur au minimum requis et ses conditions de vie égales au minimum requis ($x_i^2 > z^2$ et $x_i^3 \geq z^3$), cet individu est pauvre selon l'approche basée sur l'union (voir équation (2.50)) car ses capacités en terme de santé sont inférieures au minimum requis au niveau de la santé ($x_i^1 < z^1$).

Axiome 2.3.5 La focalisation faible (WF) : L'axiome de focalisation faible considère que la mesure de pauvreté est *uniquement* indépendante des quantités d'attributs des individus non pauvres. Ainsi, on ne tient pas compte des attributs des individus non pauvres comme dans la focalisation forte. Formellement, soit $X = [x_i^j]$ et $\tilde{X} = [y_i^j]$, s'il existe $\mathbf{x}_i \in \Omega$, $x_i^p \geq z^p$, avec $p \in \{1, \dots, m\}$ et (a) pour tout $j \in \{1, \dots, m\}$, $y_i^j = x_i^j + \varepsilon$ (avec $\varepsilon > 0$), (b) $x_i^q = y_i^q$ pour tout $q \neq j$, and (c) $x_k^l = y_k^l$ pour tout $k \neq i$ et pour tout $l \in \{1, \dots, m\}$, $i, k \in \{1, \dots, n\}$ alors $\mathbf{P}(\mathbf{z}, \tilde{X}) = \mathbf{P}(\mathbf{z}, X)$.

Contrairement à la focalisation *forte*, la *focalisation faible* autorise la compensation (ou substituabilité) entre les attributs *maigres* et *non maigres* des individus pauvres. Par exemple, considérant que le vecteur des seuils de pauvreté soit donné par $\mathbf{z} = (0.8; 0.5; 0.6)$ et le vecteur des attributs d'un individu $\mathbf{x}_i = (0.7; 0.95; 0.6)$. Cet individu est non pauvre selon la focalisation faible. L'axiome de focalisation faible est avantageux lorsque les besoins fondamentaux sont échangeables et portent sur les attributs de même type. Ce principe pose des difficultés analytiques non négligeables (à cause de la mixité des attributs) qui amènent les auteurs à opter uniquement pour l'axiome de la focalisation forte.

Axiome 2.3.6 La monotonie (MN) : L'axiome de monotonie impose que la mesure de pauvreté ne devrait pas augmenter (ou diminuer) lorsqu'un individu pauvre sur un attribut améliore sa situation par rapport à cet attribut. Formellement, soit $X = [x_i^j]$, $\tilde{X} = [y_i^j]$, $\mathbf{z} \in Z$ et $j \in \{1, \dots, m\}$, si (a) pour tout \mathbf{x}_i tel que $y_i^j = x_i^j + \varepsilon$, où $x_i^j < z^j$ et $\varepsilon > 0$, (b) $x_k^j = y_k^j$ for all $k \neq i$, et (c) $x_i^l = y_i^l$ pour tout $l \neq j$ et pour tout \mathbf{x}_i , $i, k \in \{1, \dots, n\}$ alors $\mathbf{P}(\mathbf{z}, \tilde{X}) \leq \mathbf{P}(\mathbf{z}, X)$.

Axiome 2.3.7 Le principe de la population (PP) : Le principe de la population exige de construire des mesures de pauvreté insensibles à la taille de la population. Formellement, soit $X = [x_i^j]$, $\mathbf{z} \in Z$ et $k \in \mathbb{N}$, $\mathbf{P}(\mathbf{z}, X^k) = \mathbf{P}(\mathbf{z}, X)$, où X^k est la k^e réplique de X .

En d'autres termes, l'axiome du principe de la population préconise que la mesure de pauvreté demeure inchangée lorsque la la distribution multidimensionnelle X est répliquée k fois. Cet axiome est nécessaire pour les comparaisons ordinales de pauvreté entre pays car il implique que les distributions de besoins essentiels correspondant à des populations de tailles différentes peuvent être comparées.

Axiome 2.3.8 L'invariance aux variations d'échelle (SI) : L'axiome de l'invariance aux variations d'échelle exige que la mesure de pauvreté $\mathbf{P}(\mathbf{z}, X)$ soit invariante par rapport à des transformations des échelles de mesure des attributs et des seuils. Formellement, soit $X = [x_i^j]$ et $\mathbf{z} \in Z$, $\mathbf{P}(\mathbf{z}, X) = \mathbf{P}(\Lambda\mathbf{z}, \Lambda X)$, where $\Lambda = \text{diag}(\lambda^1, \dots, \lambda^m)$ et $\lambda^j > 0$ for all $j \in \{1, \dots, m\}$.

Cet axiome permet de conclure que la fonction de pauvreté individuelle doit vérifier l'égalité suivante :

$$\mathbf{f}_i(\mathbf{z}, X) = \Lambda \mathbf{f}_i(\mathbf{z}, X) \quad (2.66)$$

avec $\Lambda x_i = \left(\frac{x_i^1}{z^1}, \dots, \frac{x_i^j}{z^j}, \dots, \frac{x_i^m}{z^m} \right)$. D'après l'axiome de l'invariance aux variations d'échelle, la distribution multidimensionnelle X peut faire intervenir des attributs avec différentes échelles de mesure (ou unité de mesure). Ainsi, la mesure de la pauvreté en terme de CFA donnerait le même résultat en terme des Euros ou des Dollars U.S. américains. La mesure de la pauvreté doit donc être insensible aux variations d'échelle. Seule la *distance relative* entre les quantités des attributs et leurs seuils respectifs est nécessaire.

Axiome 2.3.9 La monotonie par sous-groupe (SGM) : L'axiome de la monotonie par sous-groupe recommande qu'une mesure multidimensionnelle de pauvreté soit sensible au niveau du bien-être des différents segments de la population présentant des caractéristiques homogènes, telles que l'âge, le genre, le lieu de résidence, etc. Par exemple si nous supposons une population totale décomposée en deux sous-groupes (a et b), on obtient la formulation suivante : soient les deux distributions multidimensionnelles $X = \begin{bmatrix} X_a \\ X_b \end{bmatrix}$ et $Y = \begin{bmatrix} Y_a \\ Y_b \end{bmatrix}$

$$\left[\left(\mathbf{P}(\mathbf{z}, X_a) > \mathbf{P}(\mathbf{z}, Y_a) \right) \wedge \left(\mathbf{P}(\mathbf{z}, X_b) > \mathbf{P}(\mathbf{z}, Y_b) \right) \right] \implies \mathbf{P}(\mathbf{z}, X) > \mathbf{P}(\mathbf{z}, Y) \quad (2.67)$$

$\mathbf{P}(\mathbf{z}, X)$ vérifie l'axiome SGM si

$$\mathbf{P}(\mathbf{z}, X) = F\left(\frac{1}{n}\mathbf{f}_1(\mathbf{z}, X), \dots, \frac{1}{n}\mathbf{f}_i(\mathbf{z}, X), \dots, \frac{1}{n}\mathbf{f}_n(\mathbf{z}, X)\right) \quad (2.68)$$

Axiome 2.3.10 La décomposabilité par sous-groupe (SD) : L'axiome de la décomposabilité par sous-groupe stipule que si la population totale est partitionnée en divers sous-groupes par rapport à une caractéristique homogène (sexe, zone, âge, région, religion, etc.), alors la mesure multidimensionnelle de pauvreté est une moyenne pondérée des mesures multidimensionnelles de pauvreté des sous-groupes. Formellement, pour tout X_1, \dots, X_K telle que $X = X_1 \cup \dots \cup X_K$ et $\mathbf{z} \in Z$,

$$\mathbf{P}(\mathbf{z}; X_1 \cup \dots \cup X_K) = \sum_{i=1}^K \frac{n_i}{n} \mathbf{P}(\mathbf{z}, X_i) \quad (2.69)$$

où pour tout i , $n_i = \text{card}(X_i)$ et $n = \sum n_i$.

Une mesure décomposable en sous-groupes permet de calculer les contributions de chaque sous-groupe à la mesure multidimensionnelle de pauvreté et facilite l'identification des groupes sociaux les plus vulnérables. Par exemple, $\frac{n_k}{n} \cdot \mathbf{P}(\mathbf{z}; X_k)$ est la contribution du $k^{\text{ème}}$ sous-groupe à l'indice de pauvreté globale tandis que $\frac{n_k}{n} \cdot \frac{\mathbf{P}(\mathbf{z}; X_k)}{\mathbf{P}(\mathbf{z}; X)}$ est la proportion en termes de contribution du $k^{\text{ème}}$ sous-groupe à l'indice de pauvreté global. L'axiome de la décomposabilité par sous-groupe est donc particulièrement important pour le ciblage des sous-groupes sensibles à la pauvreté. Ce qui est extrêmement utile pour l'amélioration de l'efficacité des politiques de réduction de la pauvreté.

Définition 2.3.5 Le transfert progressif de Pigou-Dalton (PDPT) : $X = [x_i^j]$ est dit être obtenue à partir de $\tilde{X} = [y_i^j]$ par le transfert progressif de Pigou-Dalton de l'attribut j d'une personne pauvre s'il existe $\mathbf{x}_i, \mathbf{x}_k$: (a) $y_k^j < y_i^j < z^j$, (b) $x_k^j - y_k^j = y_i^j - x_i^j > 0$, avec $y_i^j - x_i^j > 0$ et $x_k^j \leq x_i^j$, (c) $x_l^j = y_l^j$ pour tout $l \neq k$, et (d) $x_l^p = y_l^p$ pour tout $p \neq j, p, j \in \{1, \dots, m\}$ et $l, k \in \{1, \dots, n\}$.

Axiome 2.3.11 Le principe de transfert multidimensionnel (MTP) : Le principe de transfert multidimensionnel est une extension de l'assertion de Sen [1976] selon laquelle les mesures monétaires doivent être sensibles à l'inégalité des revenus parmi les pauvres. Il stipule que la mesure multidimensionnelle de pauvreté diminue (ou n'augmente pas) lorsque l'inégalité entre les attributs des pauvres diminue. Alors, pour tout $X = [x_i^j]$ et $\mathbf{z} \in Z$, si \tilde{X} est obtenue à partir de X telle que $\tilde{X} = \Pi_\sigma X$ alors $\mathbf{P}(\mathbf{z}; \tilde{X}) \leq \mathbf{P}(\mathbf{z}; X)$.

Notons que, \tilde{X} est une distribution multidimensionnelle obtenue à partir de X simplement en redistribuant les attributs des pauvres par une matrice bistochastique Π_σ .

Axiome 2.3.12 Le principe de transfert à une dimension (ODTP) : Le principe de transfert à une dimension est une forme du principe de transfert multidimensionnel qui se réfère à un transfert d'un seul attribut X_j entre deux personnes pauvres. Comme précédemment, ce principe pose qu'un transfert de Pigou-Dalton progressif (régressif) entre deux personnes pauvres ne devrait pas faire augmenter (baisser) la pauvreté. Formellement, pour tout $\tilde{X} = [y_i^j]$, si X est obtenue à partir de \tilde{X} par un transfert progressif de Pigou-Dalton d'attribut entre deux personnes pauvres, alors $P(\mathbf{z}, X) \leq P(\mathbf{z}, \tilde{X})$, où $\mathbf{z} \in Z$ est arbitraire.

Axiome 2.3.13 La non décroissance de la pauvreté suite à un réarrangement augmentant la corrélation (NDCIS) : Cette axiome est une extension de la proposition faite par [Atkinson et Bourguignon \[1982\]](#) pour mettre en exergue la situation où les réarrangements des attributs entre les individus augmentent la corrélation entre les attributs. Ces réarrangements interviennent lors d'une série de transferts d'attributs au sein de la population pauvre qui augmentent la corrélation entre les attributs sans améliorer la situation des pauvres. Après des réarrangements de ce type, sous l'hypothèse implicite qu'aucun individu ne sorte de la pauvreté, le fait que la situation d'un individu pauvre s'empire sur les différents attributs entraîne que la pauvreté ne devrait pas baisser [[Chakravarty et al., 1998](#); [Tsui, 2002](#)]. Formellement, soit X_a une sous-matrice de la distribution multidimensionnelle X correspondant aux individus en situation de pauvreté.

$$\forall \mathbf{z}, X, \quad \left[\tilde{X} \leftarrow \Pi_\sigma X_a \right] \implies P(\mathbf{z}, \tilde{X}) \geq P(\mathbf{z}, X) \quad (2.70)$$

où Π_σ est une matrice stochastique telle que $\Pi_\sigma X_a$ soit différent de la matrice de permutation des ligne de la sous-matrice X_a .

Cet axiome est valable pour les attributs substituables [[Bourguignon et Chakravarty, 2003](#)]. Notons que la substituabilité se définit comme la proximité dans la nature des attributs. En effet, lorsque les attributs sont *substituables*, un réarrangement faisant augmenter leur inter-corrélation ne devrait pas diminuer la mesure de pauvreté puisque les individus les plus pauvres sont incapables de compenser les attributs les plus maigres par les attributs les plus forts. A l'opposé lorsque les attributs sont *complémentaires*, un réarrangement faisant augmenter la corrélation des attributs ne peut pas faire augmenter la pauvreté. D'où la formulation d'un axiome opposé à l'axiome NDCIS indiqué par "*la pauvreté non croissante dans le cas d'un réarrangement qui entraîne l'augmentation de la corrélation des attributs (Non increasing poverty under correlation increasing arrangement)*". Une discussion détaillée sur les approches axiomatiques peut être trouvée dans les études [Zheng \[1997\]](#); [Chakravarty et al. \[1998\]](#); [Tsui \[2002\]](#); [Bourguignon et Chakravarty \[2003\]](#); [Fusco \[2005\]](#); [Sami et Lahga \[2008\]](#). Toutefois, définir un axiome

Measurements	Axiomatic Characterisation	
	Satisfies	Violates
Incidence de la pauvreté FGT_0	-	MN, ODTP
La profondeur de la pauvreté FGT_1	MN	ODTP
L'indice de pauvreté de Sen $\mathcal{S}(\mathbb{M}, \xi)$	FO, MN and ODTP	-
L'indice de Watts $\mathcal{W}(\mathbb{M}, \xi)$	FO, MN, ODTP and SD	-
Chakravarty <i>et al.</i> [1998]	FO, MN, SD, SI, MTP and CN	-
Bourguignon et Chakravarty [2003]	MTP, NDCIS or NICIS	-
Bourguignon et Chakravarty [2003]	MN, CN, SD and OTP or MTP	-
Bourguignon et Chakravarty [2003]	SD	-

TABLE 2.1 – Liste des mesures caractérisées axiomatiquement

sur une mesure agrégée n'est pas toujours une tâche facile parce que les propriétés désirées peuvent différer selon que nous voulons agréger les préférences ou les informations.

2.3.4.2 Liste des mesures caractérisées axiomatiquement

Sen [1976] est le pionnier de l'approche axiomatique pour la mesure de la pauvreté. L'idée d'utiliser les axiomes est basée sur l'observation que, malgré les limites de l'incidence de la pauvreté FGT_0 (voir sous-section 3.6) qui ne vérifie pas les deux axiomes de monotonie et de transfert), elle continue d'être très largement utilisée. Il en est de même de la profondeur de la pauvreté FGT_1 (utilisé par United States Social Security Administration ; see Sen [1976]) qui vérifie l'axiome de monotonie mais ne vérifie pas l'axiome de transfert. Ainsi, Sen a élaboré une nouvelle mesure de la pauvreté au vue de la violation de ces conditions élémentaires par les mesures existantes sur la pauvreté. La liste de quelques mesures de pauvreté caractérisées axiomatiquement est donnée par le tableau suivant : De manière générale, une mesure multidimensionnelle de pauvreté $P(z, X)$ doit satisfaire SM, PP, MN, FO et CN [see Tsui, 2002].

2.3.5 L'approche par les capacités

Les origines de l'approche par les capacités remontent à la série des publications de Sen sur les critiques des approches traditionnelles de l'économie de bien-être écrits dans le début des années 80 [see Sen, 1979a, 1985, 1993] où il développa les concepts de *capabilité* et *fonctionnement* pour évaluer le bien-être des individus. Les fonctionnements se réfèrent aux accomplissements d'un individu c'est-à-dire ce que il ou elle aspire *faire* ou *être*. La capabilité, qui prend en compte la notion de liberté, se réfèrent à l'aptitude ou l'opportunité réelle d'un individu

dans l'accomplissement d'un fonctionnement donné.

Sen [1985] fût le premier à formaliser le concept de capabilité. Soit \mathbf{x}_i le vecteur des commodités possédées par un individu (ou un ménage) $i \in \{1, \dots, n\}$. Soit $\psi(\cdot)$ la fonction de conversion d'un vecteur de commodités en vecteur caractéristique. Soit $g_i[\cdot]$ la fonction d'utilisation personnelle d'un individu i reflétant ce qu'il peut actuellement accomplir. Soit \mathcal{L}_i l'ensemble des fonctions d'utilisation parmi lesquelles l'individu i peut faire son choix et soit $\mathcal{V}_i[\cdot]$ la fonction d'évaluation du vecteur de fonctionnements h_i qui évalue le niveau de bien-être de l'individu i . Ainsi, le vecteur de fonctionnements atteignables h_i peut être donné par l'équation (2.71) et représente les fonctionnements atteignables par l'individu i lorsqu'il choisit la fonction d'utilisation g_i pour un vecteur de commodités \mathbf{x}_i .

$$h_i = g_i \circ \psi(\mathbf{x}_i) = g_i[\psi(\mathbf{x}_i)] \quad (2.71)$$

Le vecteur h_i représente ce que l'individu i est capable de faire et d'être avec son vecteur de commodités \mathbf{x}_i . Le bien-être de l'individu i peut alors être donné par l'équation (2.72).

$$\tilde{\mathcal{V}}_i = \mathcal{V}_i[h_i] = \mathcal{V}_i[g_i \circ \psi(\mathbf{x}_i)] \quad (2.72)$$

Par l'utilisation de $\mathcal{V}_i[\cdot]$, il est possible de caractériser les évaluations du bien-être que l'individu i peut potentiellement atteindre comme le montre l'équation (2.73) ci-dessous.

$$\tilde{\mathcal{V}}_i = \left\{ \mathcal{V}_i : \exists \mathcal{V}_i[h_i], h_i \in \mathcal{Q}_i(\mathbf{x}_i) \right\} \quad (2.73)$$

Delà, on peut dériver un ensemble de vecteurs de fonctionnements atteignables $\mathcal{R}_i(\mathbf{x}_i)$ donné par l'équation (2.74).

$$\mathcal{R}_i(\mathbf{x}_i) = \left\{ h_i : h_i = g_i \circ \psi(\mathbf{x}_i), \forall g_i[\cdot] \in \mathcal{L}_i \right\} \quad (2.74)$$

Donc, le vecteur de fonctionnements atteignables devient :

$$\mathcal{Q}_i(\mathbf{x}_i) = \left\{ h_i : h_i = g_i \circ \psi(\mathbf{x}_i), \forall g_i[\cdot] \in \mathcal{L}_i \wedge \forall \mathbf{x}_i \in \tilde{\mathbb{X}}_i \right\} \quad (2.75)$$

et, $\mathcal{Q}_i(\mathbf{x}_i)$ représente formellement l'ensemble des capabilités de l'individu i , c'est-à-dire la liberté que l'individu i a de choisir parmi tous les fonctionnements alternatifs possibles selon ses caractéristiques personnelles et son environnement social.

Les critiques portant sur l'approche par les capabilités de Sen se focalisent sur l'ambiguïté de la théorie des capabilités et les questions liées à l'agrégation en un seul indice des multiples dimensions de capabilités. Bénicourt [2004, 2006] évoque l'impasse de l'approche portant sur plusieurs critères éthiques. Navarro [2000] met

en lumière le manque d'analyse des relations de pouvoirs entre les organisations internationales et les pays en développement. Pogge [2002] défend plutôt l'approche de Rawls des biens premiers contre l'approche par les capacités. Nussbaum [1987] reproche à Sen de n'avoir pas fourni une liste claire de capacités, et de là [Nussbaum, 2000, 2003, 2005] a isolé une liste de capacités pouvant être considérée comme centrale dans la vie humaine. Toutefois, malgré les défauts de l'approche par les capacités, elle est plus complète (en terme de conceptualisation de la pauvreté) que les autres approches existantes.

2.3.6 Les nouveaux indices multidimensionnels de pauvreté

2.3.6.1 Approche de Alkire-Foster

Alkire et Foster [2011] ont proposé une nouvelle méthodologie de mesure multidimensionnelle de la pauvreté consistant en une méthode d'identification ρ_k qui étend les approches traditionnelles d'intercession et d'union et la classe des mesures de la pauvreté FGT_α . Ils ont utilisé deux formes de seuils à l'étape d'identification : un dans chaque dimension pour déterminer si une personne est déprivée dans cette dimension, et une seconde à travers les dimensions qui identifient les pauvres en comptant les dimensions dans lesquelles une personne est déprivée. À l'étape d'agrégation, ils ont utilisé les mesures de FGT, convenablement ajustées pour prendre en compte la multidimensionnalité. En leur papier [Alkire et Foster, 2011], ils présentent un certain nombre d'axiomes et ils prouvent que leur méthodologie satisfait une gamme des propriétés souhaitables y compris la décomposabilité. L'approche de Alkire-Foster nous permet de mesurer la pauvreté au niveau individuel (individu ou ménage) en tenant compte des critères multiples (dimensions et indicateurs). Leur approche est souple et peut être employée avec différentes dimensions et indicateurs pour créer des mesures d'incidence, d'intensité et de profondeur de la pauvreté, aussi bien que l'inégalité parmi les pauvres à différentes sociétés et situations selon le type de données disponibles. Se référer à Alkire et Foster [2011] pour plus de détails concernant la méthodologie et quelques applications de l'approche aux Etats-Unis et en Indonésie.

2.3.6.2 L'indice multidimensionnel de pauvreté (MPI)

En 2010, l'OPHI (Oxford Poverty and Human Development Initiative) et le PNUD [voir UNDP, 2010] ont lancé l'indice multidimensionnel de pauvreté (MPI) comme nouvelle mesure qui donne une image *multidimensionnelle* des personnes vivantes dans la pauvreté. Le MPI a été créé par Alkire et Santos [2010] en utilisant l'approche de Alkire-Foster [voir Alkire et Foster, 2011]. Selon Alkire et Santos [2010], le MPI est un indice de la pauvreté multidimensionnelle aiguë qui reflète des déprivations dans des services très rudimentaires et des fonctionnements

humains des personnes à travers 104 pays en voie de développement. Le MPI identifie des déprivations à travers trois dimensions : santé, éducation et niveau de vie (qui sont reflétées par 10 indicateurs à importance égale), et présente le nombre de personnes qui sont “*multidimensionnellement*” pauvres et les déprivations auxquelles elles font face au niveau ménage. Le MPI est calculé comme suite :

$$MPI = H \times A \quad (2.76)$$

où H est le pourcentage des personnes qui sont pauvres selon MPI (incidence de la pauvreté) et A l'intensité moyenne de la pauvreté selon MPI à travers les pauvres (%). Le MPI est plus approprié pour pays moins développés. Elle permet de capturer l'étendue des déprivations dans les pays en voie de développement comme en Asie du Sud et en Afrique Sub-Saharienne, et dans les pays latino-américains les plus pauvres. Le MPI montre aussi l'ampleur de pauvreté au delà des mesures monétaires. En bref, le MPI nous permet de calculer et de donner un aperçu du recouvrement des déprivations selon le consensus international et en conformité avec les objectifs millénaires du développement [voir [UNDP, 2010](#); [Alkire et Santos, 2010](#)].

Notons que, l'indice MPI utilise les données niveau-ménage, puis agrège ceux-ci au niveau-pays. Comme nous l'avons mentionné plus haut, l'indice MPI identifie trois grandes dimensions – santé, éducation et niveau de vie – de même poids pour construire l'index composite multidimensionnel de pauvreté. Cette façon de faire pose le problème du choix des poids et de la façon dont ces derniers doivent être agrégés. Le choix des “poids” est en soi un problème de décision et il est crucial d'*éliciter* à travers des tests de robustesse. L'influence négative des effets de compensation ne sont ni étudiées ni considérées dans l'index MPI.

Le but de ce chapitre était de fournir une étude plus récente de la littérature sur les mesures de la pauvreté multidimensionnelle. En outre, nous avons présenté certains inconvénients sur les mesures existantes. Deux des mesures de la pauvreté multidimensionnelle ont particulièrement attiré notre attention : L'approche par les capacités de Sen [[Sen, 1985, 1993](#)] et le MPI créé par [Alkire et Santos \[2010\]](#). Les deux mesures ont été développés en considérant la nature multidimensionnelle de la la pauvreté et les deux approches présentent également des inconvénients. Toutefois, de nombreux auteurs [tels que [Saith, 2001](#); [Robeyns, 2009, 2010](#); [Nussbaum, 2000, 2003, 2005](#); [Alkire et Santos, 2010](#); [Sen, 1985, 1993](#)] et organisations internationales (telles que le PNUD [[UNDP, 2010](#)] et la banque mondiale [[World Bank, 2001](#)]) s'accordent pour dire que l'approche par les capacités est plus complète et plus appropriée pour évaluer le bien-être des personnes.

Les capacités intègrent explicitement la ‘liberté’ de bien-être (ou la liberté d'atteindre un certain niveau de vie) et, implicitement, la promotion de la justice

sociale. La principale pierre d'achoppement de cette approche est sa difficulté à être opérationnalisée en raison de la subjectivité de la pauvreté. La pauvreté est en effet multidimensionnelle et évolutive, et son mesurage ne consiste pas en une représentation d'une situation objective, mais plutôt à la construction d'un instrument de poursuite d'une certaine politique. Le mesurage de la pauvreté va au delà d'une simple identification des personnes vivant en dessous du seuil de pauvreté (\$ 1,25 par jour) ou de l'agrégation de plusieurs indicateurs en un composite unique. Le problème n'est pas seulement de savoir si quelqu'un est pauvre, mais de savoir ce que nous pouvons faire pour lui permettre de ne pas être pauvre dans l'avenir. Cette idée est le point de départ à partir duquel nous essayerons de trouver quelques directions opérationnelles en utilisant la théorie par les capacités et l'aide à la décision pour le mesurage de la pauvreté. Notre postulat souligne la nécessité de considérer le problème de mesurage de la pauvreté comme un problème de décision et d'aborder la question avec cela à l'esprit.

Deuxième partie

Construction d'indices : méthodologie générale

Méthodologie pour l'analyse et le mesurage de la pauvreté

Notre méthodologie est basée sur l'usage des techniques de mesurage ordinaire dérivées des méthodes de clustering et l'aide multicritère à la décision. L'hypothèse principale est qu'il n'existe pas de mesure "objective" de la pauvreté (un phénomène social extrêmement complexe), mais seulement une mesure opérationnelle destinée à une certaine politique de réduction de la pauvreté dans une région donnée. La question de la signification est ainsi analysée à la fois d'un point de vue théorique (méthodes de clustering et aide à la décision) et à partir d'un domaine opérationnel (politique efficace).

3.1 Concepts de base

3.1.1 Commodités, fonctionnements et capacités

Les origines de l'approche par les capacités remontent à une série d'articles scientifiques critiquant l'approche traditionnelle de l'économie du bien-être que Sen a rédigé au début des années 80 [voir Sen, 1979a, 1985, 1993] où il a développé les concepts de '*commodity*', '*functioning*', et '*capability*' pour l'évaluation du bien-être des individus. Comme l'a mentionné Alkire [2005], "*the goal of both human development and poverty reduction should be to expand the capability that people have to enjoy 'valuable beings and doings'. They should have access to the positive resources they need in order to have these capabilities. And they should be able to make choices that matter to them.*" L'approche par les capacités apparaît comme la plus influente tentative récente pour l'évaluation de l'accomplissement d'une personne ('*bien-être*'), les opportunités réelles que cette personne possède ('*avantage*') et la quantité de '*bonheur*' générée dans la vie de cette personne. Initialement, chaque individu dispose d'un ensemble de '*dotations*' (tels que les talents naturels, les habilités physiques, les compétences intellectuelles, etc.) à partir duquel il produit les commodités qui seront ensuite échangées au sein de la société. Les commodités ne sont pas nécessairement traduites en bien-être mais elles sont plutôt conceptualisées en termes de caractéristiques d'une personne. Par exemple, la possession d'un véhicule permet au propriétaire de bénéficier de toutes les proprié-

tés du véhicule, lesquelles peuvent être utilisées pour satisfaire la mobilité, pour transporter des marchandises, pour obtenir le bonheur de voyage ou pour la location. Comme l'a souligné Sen [1985], les caractéristiques des commodités ne renseignent pas sur ce qu'un individu est capable de faire avec ces propriétés. Par exemple, une personne peut avoir un véhicule et des compétences en mécanique mais, être "mal-transportée" parce qu'il n'existe aucune station service dans son village. Ce point de vue a conduit Sen à considérer les 'fonctionnements' d'un individu. Un fonctionnement est l'accomplissement d'un individu, c'est-à-dire ce qu'il réussit à 'faire' et à 'être' avec ses commodités et leurs caractéristiques en considérant ses caractéristiques personnelles, les facteurs sociaux et environnementaux au-delà de son contrôle. Par exemple, disposant d'un véhicule personnel (commodité) pour faciliter sa mobilité (caractéristique), certains individus peuvent accomplir le fonctionnement 'adéquatement transportés', tandis que d'autres individus (ayant la même commodité et la même caractéristique), mais sans aucune station service dans leur village, peuvent accomplir le fonctionnement 'pauvrement transportés'. De ce fait, 'être adéquatement transporté' doit se distinguer du fait d' 'être propriétaire' d'un véhicule. Le sous-ensemble de la totalité de tous les vecteurs de fonctionnement représente les capacités d'un individu [see Sen, 1985]. Une capacité reflète l'habileté d'une personne à accomplir un sous-ensemble de fonctionnement qu'elle a choisi parmi la totalité des fonctionnements. Par exemple, une personne peut être capable d'acheter un véhicule pour accomplir le fonctionnement 'être bien-transportée', mais il peut choisir de ne pas l'être et continuer d'aller à pied ou à bicyclette.

Le sous-ensemble de fonctionnements compatible avec un ensemble de commodité donné (resultant de treillis de fonctionnements) sera appelé l'ensemble des 'capacités réalisables'. Certaines commodités (tels que le revenu, la taille du ménage, etc.) peuvent être repartis sur toutes les dimensions du bien-être. Ce sont les 'commodités génériques'. Les capacités réalisables peuvent être étendues en utilisant les commodités génériques qui permettent d'augmenter certains fonctionnements atteignables. Nous obtenons alors un nouvel ensemble que nous appelons l'ensemble des 'capacités étendues'. Dans le but de bien comprendre les sections suivantes, nous introduisons les définitions suivantes :

Définition 3.1.1 *un 'ménage' est une unité socio-économique de base au sein de laquelle les différents membres, apparentés ou non, vivent dans la même maison ou concession, mettent en commun leurs ressources et satisfont en commun à l'essentiel de leurs besoins alimentaires et autres besoins essentiels non alimentaires, sous l'autorité de l'un d'entre eux appelé chef de ménage (CM).*

Définition 3.1.2 *Un 'cluster' est un sous-ensemble de ménages similaires ou semblables en termes de distribution de commodités, et les ménages des clusters différents ne sont pas similaires.*

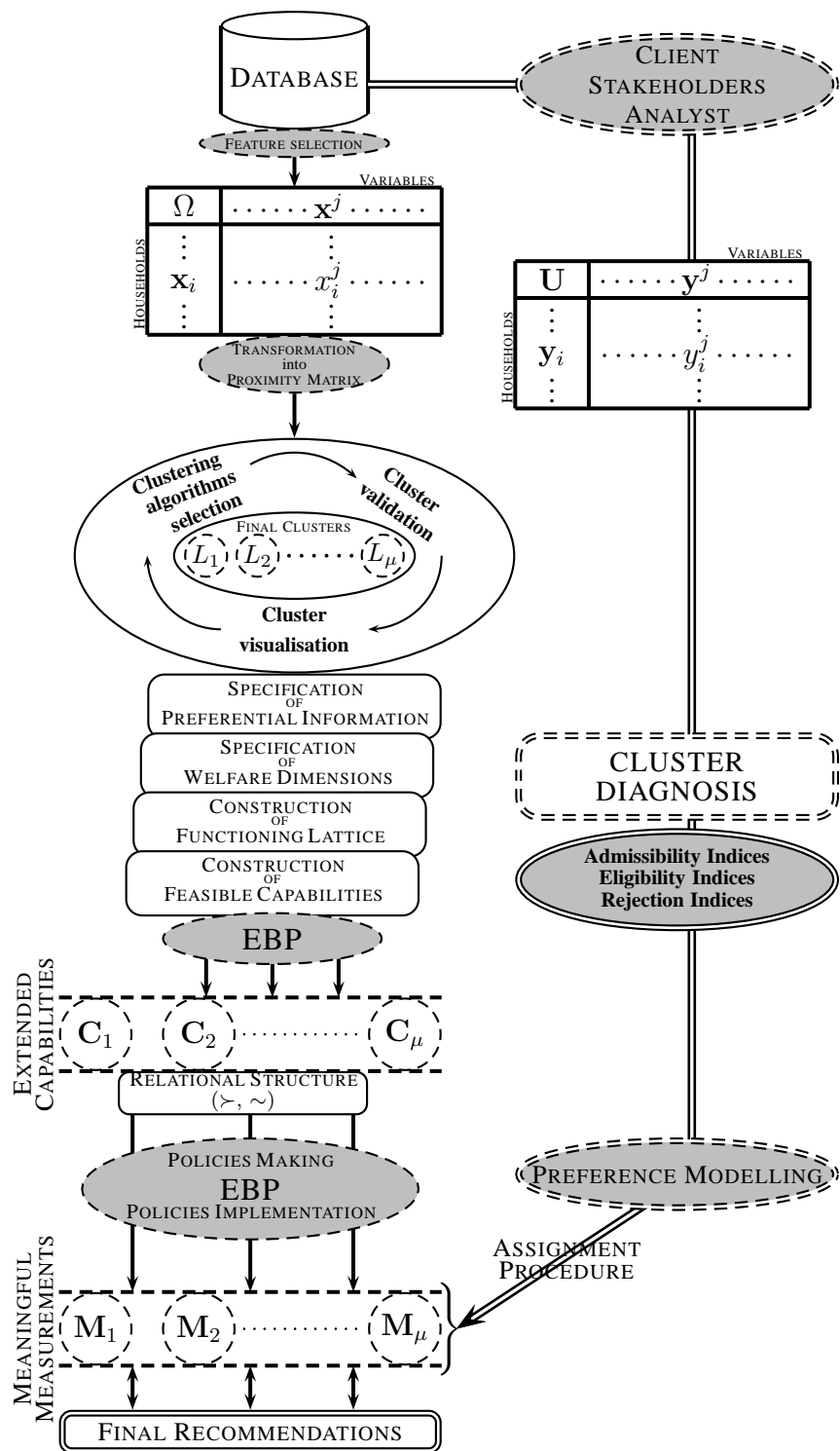


FIGURE 3.1 – Schéma général de la méthodologie MDPM

3.1.2 Notations et terminologies

L'unité d'analyse de base est un ménage. Cependant, en pratique, notre méthodologie peut être appliquée tant au niveau individuel comme une personne, et à un niveau collectif, comme un groupe ou une classe en utilisant la technique standard de l'hypothèse d'un '*individu représentatif*'. Initialement, chaque ménage \mathbf{x}_i est doté d'un ensemble de ressources. Ces ressources sont transformées en commodités (*marchandises*) qui peuvent être échangées au sein de la société. Dans cette thèse, nous n'aborderons pas la question visant à savoir comment les dotations sont converties en commodités. Nous supposons que la table de données d'entrée représente l'ensemble des vecteurs de commodités des ménages qui a été déterminé à partir d'une enquête sur les conditions de vie des ménages. Nous supposons que la table de réponses (table de données) utilisée comme données d'entrée représente l'ensemble des vecteurs des commodités des ménages qui a été obtenu à partir d'une enquête annuelle sur les conditions de vie des ménages. Ainsi, le vecteur de commodités d'un ménage peut être définie comme étant l'ensemble de tous les '*biens et services*' qu'il déclare légalement atteindre avec ses dotations.

Considérons une *distribution multidimensionnelle* d'une population Ω de n ménages donnée par la table de données $n \times m$ (3.1).

$$\Omega = \begin{matrix} & X_1 & \dots & X_j & \dots & X_m \\ \mathbf{x}_1 & \left[\begin{array}{cccccc} x_1^1 & \dots & x_1^j & \dots & x_1^m \\ \vdots & & \vdots & & \vdots \\ \mathbf{x}_i & \left[\begin{array}{cccccc} x_i^1 & \dots & x_i^j & \dots & x_i^m \\ \vdots & & \vdots & & \vdots \\ \mathbf{x}_n & \left[\begin{array}{cccccc} x_n^1 & \dots & x_n^j & \dots & x_n^m \end{array} \right] \end{array} \right] \end{array} \right] \end{matrix} \quad (3.1)$$

où x_i^j est la commodité du ménage $i \in \{1, 2, \dots, n\}$ sur son j ème attribut, $X = X_1 \times \dots \times X_m$ et $\Omega \subset X$. L'interprétation est la suivante : on considère les ménages \mathbf{x}_i , $i \in \{1, 2, \dots, n\}$, évalués sur un ensemble $\mathcal{J} = \{1, 2, \dots, m\}$ de variables. L'ensemble X_j regroupe tous les niveaux possibles réalisables ou atteignables par un ménage sur la j ème variable ($j \in \mathcal{J}$). L'ensemble X_j représente le j ème ensemble d'évaluations sur les n ménages. Le vecteur de commodités du ménage \mathbf{x}_i représente l'ensemble des évaluations du i ème ménage sur les m variables et $\mathbf{x}_i \in X$. L'ensemble X représente l'ensemble de tous les vecteurs de commodités possibles sur les m variables.

Comme il est de coutume en analyse de données, nous considérons deux caractéristiques des données : *type de données* et *échelle de données*. Les *type de données* se rapportent au degré de quantification dans les données et *l'échelle de données* indique l'importance relative des nombres. Les données d'entrée peuvent être de type *binnaire* (par exemple "oui-non"), *discrète* et *continue*. Les données

d'entrée peuvent également être définies sur des échelles *qualitatives* (nominales et ordinales) ou sur des échelles *quantitatives* (intervalle et ratio). Les valeurs possibles d'une variable qualitative (nominale ou discrète) sont appelées "modalités de la variable".

Remarque 3.1.1 Dans le cas des variables continues ou quantitatives, la moyenne arithmétique μ_j sur la j ème variable est donnée par $\mu_j = \frac{1}{n} \sum_{i=1}^n x_i^j$ et l'écart-type σ_j sur la j ème variable est donnée par $\sigma_j^2 = \frac{\sum_{i=1}^n (x_i^j - \mu_j)^2}{n-1}$.

3.2 Transformation de la table des réponses en matrice de proximité

Initialement, nous devons sélectionner les caractéristiques (*sélection des individus* et *sélection de variables*) sur lesquels la classification non supervisée ('clustering' ou 'segmentation') doit être effectuée de manière à représenter l'information autant que possible de manière conforme à notre intérêt. Selon la structure de la table de réponses, il peut s'avérer nécessaire de standardiser certaines variables avant de calculer la matrice de proximité.

3.2.1 Standardisation des variables

Parfois, il est utile de supprimer les effets du choix d'origine et d'échelle dans la mesure des variables à travers la standardisation des variables. [Milligan et Cooper \[1988\]](#) soutiennent qu'un problème méthodologique lors de l'application des techniques de classification consiste à savoir s'il faut standardiser ou non les variables d'entrée avant de procéder au calcul d'une mesure de dissimilarité. La standardisation des variables est surtout recommandée dans les cas où la mesure de dissimilarité, telle que la distance euclidienne, est sensible aux différences dans les grandeurs ou les échelles des variables d'entrée. Elle permet ainsi d'ajuster l'ampleur des scores et la pondération relative des variables [voir [Anderberg, 1973](#); [Milligan et Cooper, 1988](#), pour les applications]. Pour des raisons empiriques, nous avons considéré deux procédures de standardisation des variables : *standardisation z-score* et *standardisation variable range*.

La "standardisation z-score" transforme chaque variable continue ou quantitative en une variable de moyenne zero et de variance 1 comme suite :

$$\Omega[\dots, \mathbf{x}^j, \dots] := \frac{(\mathbf{x}^j - \mu_j)}{\sigma_j} \tag{3.2}$$

La “standardisation variable ranges” transforme chaque variable ordinale en une valeur entre 0 (valeur minimale) et 1 (valeur maximale).

$$\Omega[\dots, \mathbf{x}^j, \dots] := \frac{(\mathbf{x}^j - \min\{\mathbf{x}^j\})}{(\max\{\mathbf{x}^j\} - \min\{\mathbf{x}^j\})} \quad (3.3)$$

3.2.2 Détermination de la matrice de proximité

Soit \mathbf{x}_i le i ème ménage appartenant à un échantillon $\Omega = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ de n ménages sous-ensemble de l'univers \mathbf{U} (la population totale). Les méthodes de clustering exigent qu'une mesure de proximité (d'affinité ou de ressemblance) soit établie entre des paires de ménages. De ce fait, cette étape vise à transformer la table des données en une *matrice de proximité*. Une matrice de proximité $[\delta(\mathbf{x}_i, \mathbf{x}_k)]$ accumule des paires d'indices de similitude (indices de dissimilarité) dans une matrice dans laquelle chaque ligne et colonne représente un ménage. Plus le i ème et k ème ménage *se ressemblent*, plus petit est l'indice de dissimilarité et plus grand est l'indice de similarité. On supposera ici que l'indice de proximité est une mesure de la dissimilarité $\delta(\mathbf{x}_i, \mathbf{x}_k)$ entre les i ème et k ème ménage s'il satisfait les trois conditions suivantes pour tous $\mathbf{x}_i, \mathbf{x}_k \in \Omega$:

$$(i) \delta(\mathbf{x}_i, \mathbf{x}_k) \geq 0; \quad (ii) \delta(\mathbf{x}_i, \mathbf{x}_i) = 0; \quad (iii) \delta(\mathbf{x}_i, \mathbf{x}_k) = \delta(\mathbf{x}_k, \mathbf{x}_i). \quad (3.4)$$

Notons que la troisième condition (iii) évoque que les indices de proximité sont symétriques et, prise en conjonction avec la condition (ii), cela implique qu'une matrice de proximité est une matrice triangulaire inférieure à $n(n - 1)/2$ valeurs.

Soit \mathcal{J}_1 l'ensemble des variables quantitatives, \mathcal{J}_2 l'ensemble des variables ordinales et \mathcal{J}_3 l'ensemble des variables nominales tel que $\mathcal{J} = \mathcal{J}_1 \cup \mathcal{J}_2 \cup \mathcal{J}_3$ et $\mathcal{J}_1 \cap \mathcal{J}_2 \cap \mathcal{J}_3 = \emptyset$. Pour toute partie non vide \mathcal{J}_1 (resp. \mathcal{J}_2 et \mathcal{J}_3) de l'ensemble des variables \mathcal{J} , nous désignons par $X_{\mathcal{J}_1}$ (resp. $X_{\mathcal{J}_2}$ et $X_{\mathcal{J}_3}$) la “table de données en bloc” $X_{\mathcal{J}_1} = \prod_{j \in \mathcal{J}_1} X_j$ (resp. $X_{\mathcal{J}_2} = \prod_{j \in \mathcal{J}_2} X_j$ et $X_{\mathcal{J}_3} = \prod_{j \in \mathcal{J}_3} X_j$).

Définition 3.2.1 une “table des données en bloc” $X_{\mathcal{J}_1}$ (resp. $X_{\mathcal{J}_2}$ ou $X_{\mathcal{J}_3}$) est la sous-table des données de X représentant tous les vecteurs de commodités possibles des évaluations sur les sous-ensembles de variables \mathcal{J}_1 (resp. \mathcal{J}_2 et \mathcal{J}_3).

Pour le sous-ensemble $X_{\mathcal{J}_1}$, nous utilisons la distance euclidienne pour calculer un indice de dissimilarité entre deux ménages :

$$\delta_1(x_i, x_k) = \frac{\left(\sum_{j \in \mathcal{J}_1} w_{ikj} (x_i^j - x_k^j)^2 \right)^{1/2}}{\sum_{j \in \mathcal{J}_1} w_{ikj}} \quad (3.5)$$

w_{ikj} étant égal à 1 ou 0, selon que la comparaison est valide ou non sur la j ème variable.

Pour le sous-ensemble $X_{\mathcal{J}_2}$, nous calculons un indice de dissimilarité en utilisant l'indice de dissimilarité de Gower's [see [Gower, 1971](#)] donnée par l'équation suivante :

$$\delta_2(x_i, x_k) = \frac{\sum_{j \in \mathcal{J}_2} w_{ikj} S_{ikj}}{\sum_{j \in \mathcal{J}_2} w_{ikj}} \quad \text{avec } S_{ikj} = |x_i^j - x_k^j| \quad (3.6)$$

Pour le sous-ensemble $X_{\mathcal{J}_3}$, nous calculons un indice de dissimilarité en utilisant l'indice de dissimilarité de Gower's [see [Gower, 1971](#)] appliqué sur des échelles nominales :

$$\delta_3(x_i, x_k) = 1 - \frac{\sum_{j \in \mathcal{J}_3} w_{ikj} S_{ikj}}{\sum_{j \in \mathcal{J}_3} w_{ikj}} \quad \text{avec } S_{ikj} = \begin{cases} 1 & \text{if } x_i^j = x_k^j \\ 0 & \text{if } x_i^j \neq x_k^j \end{cases} \quad (3.7)$$

S_{ikj} est la contribution de la j ème variable à la mesure de la similarité. Combinant les équations (3.5), (3.6) et (3.7), nous obtenons la matrice de proximité $[\Delta(x_i, x_k)]_\Omega$ qui est donnée par l'équation suivante :

$$\Delta(x_i, x_k) = \left[\sum_{l=1}^3 \delta_l(x_i, x_k) \right]^{1/\varepsilon} \quad \text{avec } \varepsilon \geq 1 \quad (3.8)$$

où $\Delta(x_i, x_k)$ est l'indice de dissimilarité cumulé entre deux ménages x_i et x_k défini sur $\mathcal{J} = \mathcal{J}_1 \cup \mathcal{J}_2 \cup \mathcal{J}_3$. Notons que, chaque δ_l est l'indice de dissimilarité défini sur \mathcal{J}_l et ε est un paramètre de sensibilité. Notons aussi que, cette façon de calculer une matrice de proximité a été adoptée afin d'éviter le codage des données. Cependant, il est possible de transformer la table de données grâce aux *techniques de codage* [voir [Diday et al., 1982](#)] et étudier les effets de ces transformations sur la recommandation finale.

3.3 La classification

La classification [voir [Diday et al., 1982](#); [Jain et Dubes, 1988](#); [Berkhin, 2002](#), pour plus de détails] est utilisée dans le but d'organiser les données en clusters de telle manière que chaque cluster se compose de ménages qui sont *similaires* entre eux en terme de distribution de commodités et *dissimilaires* aux ménages d'autres clusters.

3.3.1 Sélection d'algorithmes de classification

La littérature [[Jain et Dubes, 1988](#); [Jain et al., 1999](#); [Berkhin, 2002](#)] en classification non supervisée propose plusieurs méthodes et algorithmes de classification

parmi lesquels on peut faire un choix. Traditionnellement, les méthodes et algorithmes de classification sont globalement divisés en deux grands groupes : *hiérarchique* et *partitionnement*. Les méthodes de classification hiérarchiques transforment une matrice de proximité en une séquence imbriquée de clusters, tandis que les méthodes de classification par partitionnement génèrent une seule partition de la table de données par une découverte des clusters “naturels” présents dans les données. Les algorithmes de classification hiérarchique produisent une série imbriquée de partitions basée sur un critère de fusion ou fractionnement des clusters sur la base de la dissemblance. Les algorithmes de classification par partitionnement identifient la partition qui optimise (généralement localement) un critère de classification. Les méthodes de classification hiérarchiques exigent en entrée une matrice de proximité entre les objets et permettent à un analyste de données de visualiser comment les objets sont fusionnés en clusters, tandis que les méthodes de classification par partitionnement peuvent prendre en entrée une matrice de proximité ou des données brutes présentées sous la forme d'un tableau de données dont les variables sont du même type.

Dans cette thèse, nous avons opté pour une méthode de classification hiérarchique basée sur la méthode de Ward, également connu sous le nom *variance minimale de Ward*. C'est l'une des méthodes de classification hiérarchique les plus répandues et distincte de toutes les autres car elle utilise une analyse de la variance pour évaluer les distances entre les clusters. En bref, cette méthode vise à minimiser la somme des carrés de deux clusters (hypothétiques) qui peuvent être formés à chaque étape [voir [Ward, 1963](#), pour plus de détails concernant cette méthode]. En général, cette méthode est considérée dans la littérature [voir [Jain et Dubes, 1988](#); [Gordon, 1999](#)] comme appropriée pour la classification non supervisée, cependant, elle tend à créer des clusters de petite taille.

3.3.2 Validation des clusters

Cette étape vise à décider du meilleur nombre de clusters s'adaptant au jeu de données et de discuter sur la question de la validation de clusters. Une question commune en classification est : “*combien de clusters y a-t-il dans mes données ?*”. La recherche d'une réponse à cette question a conduit à l'élaboration des procédures d'évaluation des résultats d'un algorithme de classification, connu sous l'expression *validité des clusters* (ou ‘*cluster validity*’ en anglais). La validation des clusters se réfère à un ensemble de procédures permettant d'évaluer les résultats d'une classification d'une manière quantitative. En d'autres termes, il existe trois approches pour examiner la question de validité des clusters : l'approche externe, l'approche interne et l'approche relative. L'approche externe (sur la base de *critères externes*) évalue les résultats d'un algorithme de classification en se basant sur une structure pré-déterminée imposée à un ensemble de données et devrait refléter

notre intuition sur la structure de regroupement de l'ensemble de données. L'approche interne (basée sur *critères internes*) évalue les résultats d'un algorithme de classification en termes de quantités qui impliquent les vecteurs de l'ensemble des données elles-mêmes (la matrice de proximité par exemple). L'approche relative (basée sur le *critères relatifs*) évalue la structure de la classification en la comparant à d'autres schémas de regroupement résultant du même algorithme mais avec des valeurs de paramètres différents.

En général, nous devons définir un critère d'arrêt permettant de décider du meilleur nombre de clusters dans un ensemble de données. Ce critère d'arrêt peut être défini dans un sens probabiliste ou non, selon le type d'approche choisie. La littérature propose plusieurs méthodes de clustering visant à déterminer le meilleur nombre de cluster et à examiner la validité d'une classification. Pour des raisons empiriques nous avons choisi deux d'entre eux : l'indice de [Calinski et Harabasz \[1974\]](#) et la technique de rééchantillonnage multi-échelle (*'multiscale bootstrapping technique'*) développé par [Suzuki et Shimodaira \[2006\]](#). L'indice de [Calinski et Harabasz \[1974\]](#) est l'un des plus répandu dans la littérature de la classification offrant en général le meilleur nombre de clusters [[Milligan et Cooper, 1985](#)]. La valeur maximale de cet indice indique le meilleur nombre de clusters dans les données. La technique de rééchantillonnage multi-échelle est définie dans un sens probabiliste et fournit les p -values pour la classification hiérarchique basée sur rééchantillonnage bootstrap multi-échelles. Les clusters qui sont fortement validés par les données auront de grandes p -values [voir [Jain et Dubes, 1988](#); [Gordon, 1999](#); [Halkidi et al., 2001](#); [Berkhin, 2002](#), pour plus de détails].

3.3.3 Visualisation des clusters

Nous avons utilisé le positionnement multidimensionnel (MDS, pour multidimensional scaling) pour la visualisation de la matrice de proximité. Le positionnement multidimensionnel est un ensemble de techniques statistiques connexes souvent utilisé dans la visualisation de l'information pour explorer les ressemblances ou les dissemblances dans les données. Un algorithme du positionnement multidimensionnel considère en entrée une matrice de proximité entre les objets, puis affecte un emplacement pour chaque objet dans un espace à q dimensions, où $q \geq 2$ est donné a priori. En pratique, les objets sont représentés par des points dans un espace (généralement) à deux dimensions, tels que les distances entre les points correspondent à des différences observées aussi près que possible (voir Figure 3.2). Voir [Kruskal et Wish \[1978\]](#); [Cox et Cox \[2001\]](#); [Groenen et Velden \[2004\]](#); [Borg et Groenen \[2005\]](#) pour plus de détails.

3.3.4 Matrice modal-valuée et critère modal-valué

La description des clusters vise à transformer la table de données de grande taille en un tableau récapitulatif afin d'acquérir des connaissances. Cet exercice permet d'identifier 'l'importance relative' d'une variable dans un cluster donné. À cet effet, nous introduisons deux notions fondamentales : la *matrice modal-valuée* et le *critère modal-valué*.

Définition 3.3.1 *Un critère modal-valué, noté $X_j(L_h)$, est une caractéristique permettant de décrire chaque modalité dans la variable X_j du cluster L_h .*

Définition 3.3.2 *Une matrice modal-valuée, noté \mathbf{X} , est un tableau rectangulaire $\mu \times m$ de critères modal-valués.*

La *matrice modal-valuée* peut être considérée comme un tableau récapitulatif de la grande table des données où la j ème colonne représente la variable X_j , $j \in \mathcal{J}$ et la h ème ligne indique le cluster L_h , $h \in \{1, \dots, \mu\}$. L'intersection de la h ème ligne et la j ème colonne indique le *critère modal-valué* lequel représente la description de la variable X_j dans le cluster L_h .

Formellement, soit $\mathbf{L} = \{L_1, \dots, L_\mu\}$ l'ensemble des clusters obtenus après la classification de la population Ω . Nous considérons une distribution cluster multi-dimensionnelle pour une population Ω de n ménages avec $\mu \times m$ *matrice modal-valuée* :

$$\mathbf{X} = [\mathbf{X}^1, \dots, \mathbf{X}^j, \dots, \mathbf{X}^m] = \begin{bmatrix} X_1(L_1) & \cdots & X_j(L_1) & \cdots & X_m(L_1) \\ \vdots & & \vdots & & \vdots \\ X_1(L_h) & \cdots & X_j(L_h) & \cdots & X_m(L_h) \\ \vdots & & \vdots & & \vdots \\ X_1(L_\mu) & \cdots & X_j(L_\mu) & \cdots & X_m(L_\mu) \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_h \\ \vdots \\ \mathbf{X}_\mu \end{bmatrix} \quad (3.9)$$

$X_j(L_h)$ dénote le *critère modal-valué* du cluster $h \in \{1, 2, \dots, \mu\}$ sur la variable $j \in \mathcal{J}$ et est défini comme suite :

$$X_j(L_h) = \langle (\beta_{j1}, \pi_{j1}^h); \dots; (\beta_{js_j}, \pi_{js_j}^h) \rangle \quad (3.10)$$

où $\{\beta_{jk} : k = 1, \dots, s_j\}$ représente un ensemble de modalités (ou états possibles) de L_h sur le domaine D_j de X_j , $j = 1, \dots, m$; π_{jk}^h est une mesure non négative associée à β_{jk} selon l'équation (3.11) et s_j est le nombre de valeurs actuellement prises par D_j sur X_j .

$$\pi_{jk}^h = \frac{|\{\mathbf{x}_i \in L_h : x_i^j = \beta_{jk}\}|}{|\{L_h\}|}, \quad k = 1, \dots, s_j \quad (3.11)$$

Notons que la modalité β_{jk} peut être finie ou infinie en nombre ; quantitative ou catégorielle en valeur ; et la mesure π_{jk}^h représente la probabilité, la proportion ou la fréquence de la valeur β_{jk} dans le cluster L_h . Intuitivement, π_{jk}^h représente 'l'importance relative' de la modalité β_{jk} dans le cluster L_h c'est-à-dire le nombre d'occurrence de la modalité β_{jk} dans la variable X_j du cluster L_h .

Définition 3.3.3 Un critère modal-valué $X_j(L_h)$ est un sous-ensemble flou de l'ensemble des modalités (ou états) $\{\beta_{jk} : k = 1, \dots, s_j\}$ d'un cluster L_h sur le domaine D_j de X_j , $j \in \mathcal{J}$ défini par :

$$X_j(L_h) = \{(\beta_{jk}, \pi_{jk}^h) : k = 1, \dots, s_j\} \quad (3.12)$$

où π_{jk}^h est le poids de la modalité β_{jk} sur la variable $j \in \mathcal{J}$ associé au cluster L_h , $h \in \{1 \dots, \mu\}$.

La matrice modal-valuée de X , notée \mathbf{X} , est l'ensemble de tous les sous-ensembles de X tel que $\mathbf{X} = \mathbf{X}^1 \times \dots \times \mathbf{X}^m$ et $\mathbf{X}^j = \{\beta_{jk} : k = 1, \dots, s_j\}$ dénote le sous-ensemble de modalités (ou états) sur le domaine D_j de X_j . Dans le cas particulier où $X_j \subseteq \mathbb{R}$, nous posons $\mathbf{X}^j \subseteq \mathbb{R}$. \mathbf{X}_h , appelé 'cluster modal-valué', représente la description du cluster L_h sur la matrice modal-valuée \mathbf{X} donnée par l'équation (3.13) :

$$\mathbf{X}_h = \langle X_1(L_h); \dots; X_m(L_h) \rangle \quad (3.13)$$

3.3.5 Exemple illustratif

En guise d'exemple, considérons la table des données (see Table 3.1) obtenue à partir d'une enquête sur les conditions de vie des ménages d'une population donnée. Pour cet exemple, nous considérons 12 ménages $\Omega = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8, \mathbf{x}_9, \mathbf{x}_{10}, \mathbf{x}_{11}, \mathbf{x}_{12}\}$ évalués sur 7 variables (une variable quantitative et 6 variable qualitative) :

Ménages	Variables						
	X_1	X_2	X_3	X_4	X_5	X_6	X_7
x_1	250	N	S	N	F	A	W
x_2	4500	Y	V	Y	V	N	Be
x_3	1500	Y	A	Y	F	S	U
x_4	200	N	S	N	F	A	W
x_5	800	N	S	N	F	S	W
x_6	5000	Y	V	Y	V	N	Be
x_7	2500	Y	A	Y	A	S	U
x_8	600	N	S	N	F	A	W
x_9	2000	Y	A	Y	A	S	Be
x_{10}	6500	Y	V	Y	V	N	U
x_{11}	1000	N	A	N	F	A	W
x_{12}	2800	Y	A	Y	A	S	Be

TABLE 3.1 – Exemple illustratif

X_j	Description des variables
X_1	Revenu mensuel pour les besoins de base des ménages en euros.
X_2	Y at-il une salle équipée pour la cuisine ? Oui(Y), Non(N).
X_3	Type de logement ? Villa(V), Immeuble(A), Maison Individuelle Simple(S).
X_4	Propriétaire d'un véhicule, voiture ou camion ? Oui(Y), Non(N).
X_5	Durée pour atteindre le moyen de transport public le plus proche (en minutes) ? [0; 14] =Très Proche(V), [15; 29] =Acceptablement Proche(A), [45; 59] =Éloigné(F).
X_6	A eu des problèmes pour répondre aux besoins alimentaires ? Jamais(N), Parfois(S), Toujours(A).
X_7	Situation Économique ? Pire Maintenant(W), Inchangée(U), Meilleure Maintenant(Be).

Nous avons effectué une analyse typologique hiérarchique via la technique de rééchantillonnage multi-échelles bootstrap [nombre de rééchantillonnage 1000 ; voir Suzuki et Shimodaira, 2006]) en utilisant la méthode de Ward [Ward, 1963] et matrice de dissimilarité basée sur la corrélation. Le meilleur nombre de clusters est de trois clusters. Ce résultat peut être confirmé par l'indice de Calinski et Harabasz [1974] lequel permet de comparer l'homogénéité des partition. La visualisation des clusters est possible via le dendrogramme des ménages et le positionnement multidimensionnel comme indiqué sur la Figure 3.2.

Considérons l'exemple illustratif (voir Table 3.1), nous obtenons les clusters $L_1 = \{x_1, x_4, x_5, x_8, x_{11}\}$, $L_2 = \{x_2, x_6, x_{10}\}$, et $L_3 = \{x_3, x_7, x_9, x_{12}\}$. La *distribution multidimensionnelle* pour cette population de 12 ménages est donnée

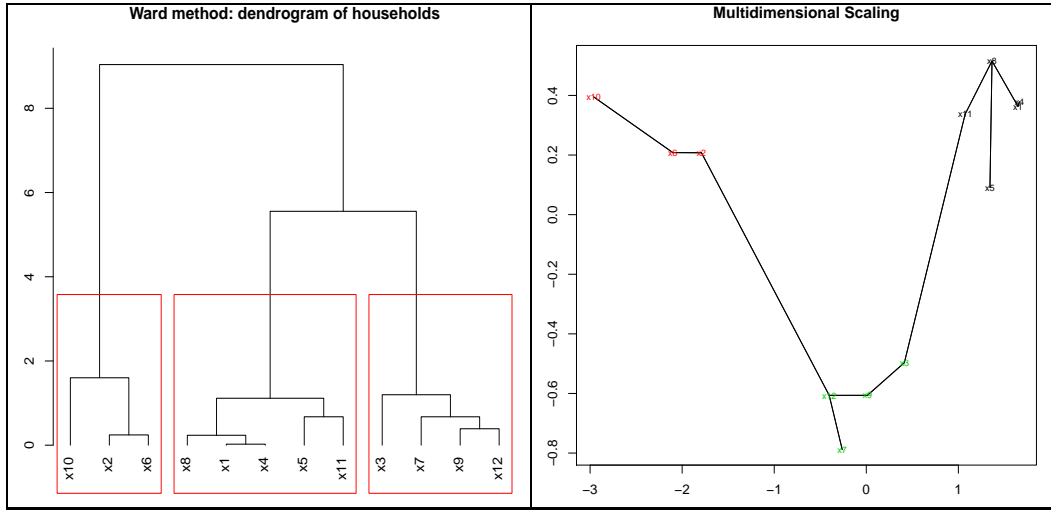


FIGURE 3.2 – Visualisation des ménages par clusters

par la table des données suivante :

$$X = \begin{bmatrix} 250.00 & N & S & N & F & A & W \\ 200.00 & N & S & N & F & A & W \\ 800.00 & N & S & N & F & S & W \\ 600.00 & N & S & N & F & A & W \\ 1000.00 & N & A & N & F & A & W \\ \hline 4500.00 & Y & V & Y & V & N & Be \\ 5000.00 & Y & V & Y & V & N & Be \\ 6500.00 & Y & V & Y & V & N & U \\ \hline 1500.00 & Y & A & Y & F & S & U \\ 2500.00 & Y & A & Y & A & S & U \\ 2000.00 & Y & A & Y & A & S & Be \\ 2800.00 & Y & A & Y & A & S & Be \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_4 \\ \mathbf{x}_5 \\ \mathbf{x}_8 \\ \mathbf{x}_{11} \\ \hline \mathbf{x}_2 \\ \mathbf{x}_6 \\ \mathbf{x}_{10} \\ \hline \mathbf{x}_3 \\ \mathbf{x}_7 \\ \mathbf{x}_9 \\ \mathbf{x}_{12} \end{bmatrix} \quad (3.14)$$

et sa matrice modal-valuée est donnée par :

$$\mathbf{X} = \begin{bmatrix} X_1(L_1) & X_2(L_1) & X_3(L_1) & X_4(L_1) & X_5(L_1) & X_6(L_1) & X_7(L_1) \\ X_1(L_2) & X_2(L_2) & X_3(L_2) & X_4(L_2) & X_5(L_2) & X_6(L_2) & X_7(L_2) \\ X_1(L_3) & X_2(L_3) & X_3(L_3) & X_4(L_3) & X_5(L_3) & X_6(L_3) & X_7(L_3) \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \end{bmatrix} \quad (3.15)$$

où chaque critère modal-valué $X_j(L_h)$ est : $X_1(L_1) = \langle ([200; 1000], 1) \rangle$, $X_2(L_1) = \langle (N, 1.00); (Y, 0.00) \rangle$, $X_3(L_1) = \langle (A, 0.20); (S, 0.80); (V, 0.00) \rangle$, $X_4(L_1) = \langle (N, 1.00); (Y, 0.00) \rangle$, $X_5(L_1) = \langle (A, 0.00); (F, 1.00); (V, 0.00) \rangle$, $X_6(L_1) = \langle (A, 0.80); (N, 0.00); (S, 0.20) \rangle$, $X_7(L_1) = \langle (Be, 0.00); (U, 0.00); (W, 1.00) \rangle$; $X_1(L_2) = \langle ([4500; 6500], 1) \rangle$, $X_2(L_2) = \langle (N, 0.00); (Y, 1.00) \rangle$, $X_3(L_2) = \langle (A, 0.00); (S, 0.00); (V, 1.00) \rangle$, $X_4(L_2) = \langle (N, 0.00); (Y, 1.00) \rangle$, $X_5(L_2) = \langle (A, 0.00); (F, 0.00); (V, 1.00) \rangle$, $X_6(L_2) = \langle (A, 0.00);$

$(N, 1.00); (S, 0.00)$, $X_7(L_2) = \langle (Be, 0.67); (U, 0.33); (W, 0.00) \rangle$; $X_1(L_3) = \langle ([1500; 2800], 1) \rangle$, $X_2(L_3) = \langle (N, 0.00); (Y, 1.00) \rangle$, $X_3(L_3) = \langle (A, 1.00); (S, 0.00); (V, 0.00) \rangle$, $X_4(L_3) = \langle (N, 0.00); (Y, 1.00) \rangle$, $X_5(L_3) = \langle (A, 0.75); (F, 0.25); (V, 0.00) \rangle$, $X_6(L_3) = \langle (A, 0.00); (N, 0.00); (S, 1.00) \rangle$, $X_7(L_3) = \langle (Be, 0.50); (U, 0.50); (W, 0.00) \rangle$.

Par exemple, $X_7(L_2)$ doit être lu comme : “considérant la variable X_7 décrivant la ‘situation économique’ d’un ménage, 67% des ménages possèdent légalement la commodité ‘Be’ et 33% la commodité ‘U’, tandis qu’aucun ménage ne possède la commodité ‘W’ dans le cluster L_2 ”.

3.4 Construction des mesures significantes

3.4.1 Caractérisation des modalités

Parfois, la fréquence π_{jk}^h de la modalité β_{jk} dans le cluster L_h , telle que définie dans l'équation (3.21), peut ne pas refléter avec exactitude l'assertion : “cette modalité est plus pertinente dans ce cluster que dans tous les autres clusters”. C'est pourquoi nous avons besoin de caractériser les modalités dans un sens probabiliste en utilisant le *test d'hypothèse*. Le test d'hypothèse est une partie essentielle de l'inférence statistique visant à déterminer la probabilité que telle ou telle hypothèse soit vraie. La caractérisation utilisée dans le présent document a été inspirée par le principe de la caractérisation statistique introduite par Morineau [1984]. L'idée consiste à vérifier si la modalité β_{jk} du critère modal-valué $X_j(L_h)$ est une caractéristique pertinente du cluster L_h au regard du fait probabilité que l'hypothèse “la modalité β_{jk} est ‘significativement’ plus abondante (ou fréquente) dans le cluster L_h que dans la population Ω ” est proche de 1. Ainsi, nous posons l'hypothèse nulle H_0 d'un tirage successif sans remise de n_h ménages parmi les n ménage de la population. H_0 assure que les fréquences π_{jk}^h et p_{jk}^h sont égales aux fluctuations aléatoires près ; où p_{jk}^h est :

$$p_{jk}^h = \frac{|\{x_i \in \Omega : x_i^j = \beta_{jk}\}|}{|\{\Omega\}|}, \quad k = 1, \dots, s_j \quad (3.16)$$

Intuitivement, p_{jk}^h représente ‘l'importance relative’ de la modalité β_{jk} dans la population c'est-à-dire le nombre d'occurrence de la modalité β_{jk} dans la variable X_j de la population Ω .

Soit N une variable aléatoire telle que $N = n_{jk}^h$ avec $n_{jk}^h = n_h \cdot \pi_{jk}^h$. Sous l'hypothèse H_0 , N suit une loi hypergéométrique, $N \sim \mathcal{H}(n, n_{jk}, n_h)$, de moyenne $E_h(N) = n_h \cdot p_{jk}^h$ et d'écart-type $\sigma_h^2(N) = n_h \cdot \frac{n - n_h}{n - 1} \cdot p_{jk}^h \cdot (1 - p_{jk}^h)$ avec $n_{jk} = n \cdot p_{jk}^h$. De là, le degré de significativité $\rho(\beta_{jk})$ sous H_0 est donné par l'équa-

tion suivante :

$$\rho(\beta_{jk}) = \text{Prob}_{H_0}\{N > \beta_{jk}\} = \text{Prob}_{H_0}\{t_h(N) > t_h(\beta_{jk})\} \quad (3.17)$$

où $t_h(N) = \frac{N - E_h(N)}{\sigma_h(N)}$ telle que $t_h(N)$ suit la loi normale centrée réduite $\mathcal{N}(0; 1)$. En pratique, une loi hypergéométrique peut être approximée par une loi normale (loi normale gaussienne, loi de Laplace-Gauss) lorsque le nombre de ménages dans le cluster est suffisamment élevé (plus de 30 ménages).

Pour chaque critère modal-valué $X_j(L_h)$ et chaque modalité β_{jk} , nous définissons l'indice de pertinence $\xi(\beta_{jk})$ par l'équation (3.18) comme suit :

$$\xi(\beta_{jk}) = \rho(\beta_{jk}) \cdot \pi_{jk}^h + \lambda_h \cdot \zeta_{jk}^h \quad (3.18)$$

où ζ_{jk}^h est donnée par l'équation (4.10) suivante et $\lambda_h = \frac{n_k}{n}$.

$$\zeta_{jk}^h = \frac{|\{x_i \in L_h : x_i^j = \beta_{jk}\}|}{|\{x_i \in \Omega : x_i^j = \beta_{jk}\}|} = \frac{n_{jk}^h}{n_{jk}}, \quad k = 1, \dots, s_j \quad (3.19)$$

À cet effet, n_{jk} (n_{jk}^h respectivement) est le nombre d'occurrence de la modalité β_{jk} dans la j ème variable de la population Ω (du cluster L_h respectivement). $\xi(\beta_{jk})$ est normalisée, avec des valeurs entre 0 et 1, en utilisant (3.20).

$$\tilde{\xi}(\beta_{jk}) = \frac{\xi(\beta_{jk})}{\max_{\{j,k\}} \{\xi(\beta_{jk})\}} \quad (3.20)$$

Définition 3.4.1 Une caractérisation modal-valuée $Y_j(L_h)$ est un sous-ensemble flou de l'ensemble des modalités (ou états) $\{\beta_{jk} : k = 1, \dots, s_j\}$ d'un cluster L_h sur le domaine D_j de X_j , $j \in \mathcal{J}$ définie par :

$$Y_j(L_h) = \{(\beta_{jk}, \xi(\beta_{jk})) : k = 1, \dots, s_j\} \quad (3.21)$$

où $\xi(\beta_{jk})$ est l'indice de pertinence de la modalité β_{jk} sur la variable $j \in \mathcal{J}$ du cluster L_h , $h \in \{1, \dots, \mu\}$.

L'indice de pertinence $\xi(\beta_{jk})$ apparaît comme un moyen de montrer sur une échelle entre 0 et 1, le niveau avec lequel l'assertion “ $\xi(\beta_{jk})$ est plus pertinent que $\xi(\beta_{j'k'})$ ” est valide. Intuitivement, $\xi(\beta_{jk})$ mesure le ‘degré de connexion’ de la modalité β_{jk} au cluster L_h considérant tous les autres clusters et la population totale Ω c'est-à-dire la ‘connexité’ entre la modalité β_{jk} dans la variable X_j et un cluster donné L_h (dans le sens l'un cause l'autre ou partage les mêmes caractéristiques que l'autre).

Remarque 3.4.1 L'indice de pertinence est une représentation numérique qui permet d'effectuer des comparaisons entre les modalités, mais il ne peut pas être traité comme une probabilité, car il ne satisfait pas la propriété “additivité”. Ces indices de pertinence sont donc utiles parce qu'ils permettent d'établir la relation entre les commodités et les niveaux de dimension du bien-être telle que nous la montreront dans les sections suivantes.

3.4.2 Spécification des informations préférentielles

Le décideur doit spécifier ses préférences conformément aux politiques, programmes ou projets qu'il envisage d'entreprendre dans une région donnée pour améliorer les conditions de vie des populations. Dans cette thèse, nous avons considéré que ce problème implique au moins cinq types d'informations préférentielles (PI) : information préférentielle concernant l'importance des variables, information préférentielle concernant les dimensions du bien-être de préférence, information préférentielle concernant les niveaux des dimensions du bien-être, information préférentielle concernant le classement des modalités et information préférentielle concernant la spécification des ménages prioritaires en terme d'intervention et des actions concrètes à entreprendre.

3.4.2.1 Les dimensions du bien-être

Le décideur en collaboration avec l'analyste doit définir l'ensemble des “*dimensions du bien-être*”. Les dimensions du bien-être permettent de spécifier et de décrire la totalité des dimensions de *fonctionnement* que le décideur souhaiterait évaluer durant le processus de décision visant à trouver une recommandation finale appropriée pour l'amélioration des conditions de vie des ménages. Ce concept est basé sur l'idée stipulant que *la façon dont les individus ou les ménages perçoivent leur position ou leur état dans une société donnée est un aspect crucial à considérer lors de l'élaboration des politiques, programmes et projets qui les concernent*. Toutefois, afin d'évaluer l'apport des politiques, nous avons besoin d'une structure analytique. C'est pourquoi nous avons introduit et développé le concept de “*dimension du bien-être*” dans cette thèse.

Définition 3.4.2 Une dimension du bien-être décrit l'ensemble de tous les fonctionnements réalisables d'un ménage sur cette dimension sous formes de niveaux du bien-être.

Formellement, soit \mathcal{J} une ensemble de variables. Pour tout sous-ensemble non vide \mathcal{J}_u de \mathcal{J} nous notons par $F_{\mathcal{J}_u}$ (respectivement, $F_{-\mathcal{J}_u}$) l'ensemble $\prod_{j \in \mathcal{J}_u} X_j$ (respectivement, $\prod_{j \notin \mathcal{J}_u} X_j$) tels que $\mathcal{J} = \mathcal{J}_1 \cup \dots \cup \mathcal{J}_p$ et $\mathcal{J}_1 \cap \dots \cap \mathcal{J}_p = \emptyset$. $F_{\mathcal{J}_u}$ est une dimension potentielle du bien-être décrivant les niveaux du bien-être sur \mathcal{J}_u et $F = \{F_{\mathcal{J}_u} : u = 1, \dots, p\}$ indique l'ensemble de toutes les dimensions potentielles du bien-être. Nous écrirons $F_{\mathcal{J}_u}(L_h)$ si ça se réfère au cluster L_h .

Remarquons que, chaque $F_{\mathcal{J}_u}$ est perçue comme l'ensemble de toutes les commodités du sous-ensemble $\prod_{j \in \mathcal{J}_u} X_j$ dans la population totale Ω et $F_{\mathcal{J}_u}(L_h)$ est perçue comme l'ensemble de toutes les commodités du sous-ensemble $\prod_{j \in \mathcal{J}_u} X_j$ dans le cluster L_h .

3.4.2.2 Les niveaux des dimensions du bien-être

Le décideur doit définir un ensemble de niveaux associés à chaque dimension du bien-être. Un *niveau d'une dimension du bien-être* permet, dans un cas plus général, de distinguer les degrés d'intensité possibles de cette dimension du bien-être. Dans le cas de la pauvreté, elle permet de distinguer le degré de nécessité des ménages en terme de politique d'intervention sur cette dimension du bien-être. Par exemple, 'Mauvais', 'Moyen' et 'Bien' peuvent être trois niveaux associés aux dimensions du bien-être 'Mobilité' ou 'Intégration Sociale'; 'Bas', 'Acceptablement Bas', 'Relativement Bas', 'Élevé' et 'Très Élevé' peuvent être cinq niveaux associés aux dimensions du bien-être 'Pollution de l'air' ou 'Risque'.

Formellement, pour la u ème dimension du bien-être $F_{\mathcal{J}_u}$, nous obtenons un ensemble de niveaux

$$F_{\mathcal{J}_u} = \{\xi_v^u : v = 1, \dots, t_u\} \quad (3.22)$$

où $u \in \{1, \dots, p\}$, ξ_v^u désigne le v ème niveau de la u ème dimension du bien-être $F_{\mathcal{J}_u}$ et t_u est le nombre de niveau(x) possible(s) actuellement pris par $F_{\mathcal{J}_u}$.

3.4.3 Construction des capacités

3.4.3.1 Détermination des indices $C_h(\mathbf{S}_{\mathcal{J}_u})$ et $C_h(\mathbf{U}_{\mathcal{J}_u})$

Dans le but de construire l'ensemble des capacités associé à chaque classe, nous introduisons deux indices : l'indice $C_h(\mathbf{S}_{\mathcal{J}_u})$ et l'indice $C_h(\mathbf{U}_{\mathcal{J}_u})$. Nous désignons par $\mathbf{M} = \{\beta_{jk} : j \in \mathcal{J} \text{ and } k \in \{1, \dots, s_j\}\}$ l'ensemble de toutes les modalités où s_j est le nombre de valeurs actuellement prises par X_j . L'idée consiste à demander directement au décideur de partitionner l'ensemble de toutes les modalités en trois sous-ensembles $\langle \mathbf{S}, \mathbf{N}, \mathbf{U} \rangle$. Nous interprétons le sous-ensemble \mathbf{S} comme contenant les modalités de Satisfaction, \mathbf{N} contient les modalités Neutre tandis que \mathbf{U} contient les modalités de Non-satisfaction. Formellement, $\mathbf{S} = \{\beta_{jk} : \beta_{jk} > \beta_{jk'} \text{ avec } k \neq k' \text{ et } k, k' \in \{1, \dots, s_j\}, j \in \mathcal{J}\}$, $\mathbf{U} = \{\beta_{jk} : \beta_{jk} < \beta_{jk'} \text{ avec } k \neq k' \text{ et } k, k' \in \{1, \dots, s_j\}, j \in \mathcal{J}\}$ et $\mathbf{N} = \{\beta_{jk} : \text{non}(\beta_{jk} > \beta_{jk'}) \text{ et } \text{non}(\beta_{jk} < \beta_{jk'}) \text{ avec } k \neq k' \text{ et } k, k' \in \{1, \dots, s_j\}, j \in \mathcal{J}\}$.

De là, pour chaque $F_{\mathcal{J}_u}$ (avec $u = 1 \dots, p$) nous posons $\mathbf{S}_{\mathcal{J}_u} = \{\beta_{jk} : j \in \mathcal{J}_u \text{ et } \beta_{jk} \in \mathbf{S}\}$ et $\mathbf{U}_{\mathcal{J}_u} = \{\beta_{jk} : j \in \mathcal{J}_u \text{ et } \beta_{jk} \in \mathbf{U}\}$. Remarquons que $\mathbf{S} = \bigcup_{1 \leq u \leq p} \mathbf{S}_{\mathcal{J}_u}$ et $\mathbf{U} = \bigcup_{1 \leq u \leq p} \mathbf{U}_{\mathcal{J}_u}$. Alors, les deux partitions $\langle \mathbf{S}, \mathbf{U} \rangle$ nous permettent d'évaluer et de ranger deux à deux les clusters en fonction d'un niveau de dimension du bien-être. L'évaluation du cluster L_h est donnée par la "force" de ses modalités de Satisfaction (voir l'équation 4.2) et la "faiblesse" de ses modalités de Non-satisfaction (voir l'équation 4.3).

$$C_h(\mathbf{S}_{\mathcal{J}_u}) = \sum_{j \in \mathcal{J}_u} w_j \left(\sum_{k \in \mathbf{S}_{\mathcal{J}_u}} \xi(\beta_{jk}) \right) \quad (3.23)$$

$$\mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u}) = \sum_{j \in \mathcal{J}_u} w_j \left(\sum_{k \in \mathbf{U}_{\mathcal{J}_u}} \xi(\beta_{jk}) \right) \quad (3.24)$$

où $h = 1, \dots, \mu$, $u = 1, \dots, p$ et w_j désigne le j ème poids positif représentant l'importance de la variable X_j . Intuitivement, $\mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u})$ (respectivement, $\mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})$) mesure le 'degré de connection' (respectivement, 'degré de déconnection') du niveau de dimension du bien-être *le plus élevé* (respectivement, *le plus bas*) de $F_{\mathcal{J}_u}(L_h)$ avec le cluster L_h considérant tous les autres et la population totale.

3.4.3.2 Les vecteurs de fonctionnements

Un 'vecteur de fonctionnements' est une collection de niveaux réalisables sur chaque dimension du bien-être. Une 'composante' du vecteur de fonctionnements est aussi appelée 'niveau d'accomplissement' sur une dimension de bien-être particulière. Formellement, nous obtenons le l ème vecteur de fonctionnements du cluster L_h par :

$$\mathbf{F}_l(L_h) = \langle F_{\mathcal{J}_1}(L_h); \dots; F_{\mathcal{J}_p}(L_h) \rangle \quad (3.25)$$

$$= \langle \xi_{\mathbf{k}_1}^1, \xi_{\mathbf{k}_2}^2, \dots, \xi_{\mathbf{k}_p}^p \rangle \quad (3.26)$$

où $\mathbf{k}_u \in \{1, \dots, t_u\}$, $u \in \{1, \dots, p\}$, $h \in \{1, \dots, \mu\}$, $l \in \{1, \dots, l_h\}$ et l_h représente le nombre de vecteurs de fonctionnements de L_h .

Remarque 3.4.2 *Trivialement, si chaque dimension du bien-être \mathcal{J}_u possède N_u niveaux de bien-être (avec $u = 1, \dots, p$), le nombre de tous les vecteurs de fonctionnements possibles N est donné par :*

$$N = N_1 \times N_2 \times \dots \times N_p \quad (3.27)$$

3.4.3.3 Le treillis des fonctionnements

Un *treillis de fonctionnements* est la visualisation de la totalité des vecteurs de fonctionnements partiellement ordonnés par une relation de dominance. Toute autre relation est établie par le client en fonction de ses préférences. En général, les vecteurs de fonctionnements sont construits indépendamment des clusters. Nous établissons un *seuil de fonctionnements* pour chaque cluster, lequel désigne le 'vecteur frontière de fonctionnements' dans le treillis de fonctionnements.

3.4.3.4 Construction des seuils de fonctionnements

Une représentation numérique évidente revient à associer un nombre réel à chaque sous-ensemble \mathbf{C}_h de l'ensemble des vecteurs de fonctionnements réalisables de telle sorte que la comparaison entre ces nombres reflète fidèlement la

relation de préférence \succsim sur les différentes dimensions de bien-être potentiels où \succ désigne la relation de préférence stricte et \sim la relation d'indifférence. Notons que, $C_a \succsim C_b$ peut être interprété comme "le niveau de vie (ou le bien-être) des ménages appartenant au cluster L_a est *au moins aussi préférable que* le niveau de vie (ou le bien-être) des ménages appartenant au cluster L_b ". De là, nous associons à chaque mesure signifiante définie sur \mathcal{J}_u et notée par $M_h(\mathcal{J}_u)$, une séquence d'alternatives qui représente concrètement les interventions appropriées aux ménages du cluster L_h . M_h est ainsi donné par l'équation suivante :

$$M_h(\mathcal{J}_u) = \vartheta_u [C_h(\mathbf{S}_{\mathcal{J}_u}), C_h(\mathbf{U}_{\mathcal{J}_u})] \quad (3.28)$$

où ϑ_u est la u ème fonction à valeur réelle permettant d'agréger $C_h(\mathbf{S}_{\mathcal{J}_u})$ et $C_h(\mathbf{U}_{\mathcal{J}_u})$. Cela conduit à un codage numérique du jugement de valeur entre $C_h(\mathbf{S}_{\mathcal{J}_u})$ et $C_h(\mathbf{U}_{\mathcal{J}_u})$ sur les différentes dimensions de bien-être potentiels $F_{\mathcal{J}_u}$, de telle manière que les sous-ensembles C_a et C_b puissent être simplement comparés en prenant la somme de ces fonctions comme suit :

$$\begin{aligned} C_a \succsim C_b &\iff \sum_{u=1}^p M_a(\mathcal{J}_u) \geq \sum_{u=1}^p M_b(\mathcal{J}_u) \\ &\iff \sum_{u=1}^p \vartheta_u [C_a(\mathbf{S}_{\mathcal{J}_u}), C_a(\mathbf{U}_{\mathcal{J}_u})] \geq \sum_{u=1}^p \vartheta_u [C_b(\mathbf{S}_{\mathcal{J}_u}), C_b(\mathbf{U}_{\mathcal{J}_u})] \end{aligned} \quad (3.29)$$

Nous construisons maintenant une *relation de jugement* $\mathcal{R}_h(\mathcal{J}_u)$ qui est un mécanisme formel permettant de relier un vecteur de fonctionnement spécifique ξ_v^u , $v = 1, \dots, t_u$ à la dimension de bien-être correspondant $F_{\mathcal{J}_u}$, $u = 1, \dots, p$. Le décideur (ou le client) doit définir les niveaux réalisables de chaque dimension du bien-être $F_{\mathcal{J}_u}$ et les seuils ε_{vu} . Nous supposons que les niveaux réalisables de chaque dimension du bien-être $F_{\mathcal{J}_u}$ sont ordonnés c'est-à-dire $\xi_1^u > \xi_2^u > \dots > \xi_{t_u}^u$, pour tout $u \in \{1, \dots, p\}$.

$$\mathcal{R}_h(\mathcal{J}_u) = \begin{cases} \xi_1^u, & \text{if } M_h(\mathcal{J}_u) \geq \varepsilon_{1u} \\ \xi_2^u, & \text{if } M_h(\mathcal{J}_u) \geq \varepsilon_{2u} \\ \vdots & \vdots \\ \xi_{t_u}^u, & \text{if } M_h(\mathcal{J}_u) \geq \varepsilon_{t_u u} \end{cases} \quad (3.31)$$

$M_h(\mathcal{J}_u)$ est donnée par l'équation (4.5) et $M_h(\varepsilon_u) \geq \varepsilon_{1u} > \varepsilon_{2u} > \dots > \varepsilon_{t_u u} \geq -M_h(\varepsilon_u)$ avec $\phi_h[\mathcal{J}_u] = \min \{C_h(\mathbf{S}_{\mathcal{J}_u}); C_h(\mathbf{U}_{\mathcal{J}_u})\}$; $\psi_h[\mathcal{J}_u] = C_h(\mathbf{S}_{\mathcal{J}_u}) + C_h(\mathbf{U}_{\mathcal{J}_u})$ et $M_h(\varepsilon_u) = 1 - \frac{\phi_h[\mathcal{J}_u]}{\psi_h[\mathcal{J}_u]}$ pour tout $h \in \{1, \dots, \mu\}$ et $u \in \{1, \dots, p\}$.

Étant donné un cluster L_h , nous pouvons établir son *vecteur sup-fonctionnements*. Nous l'appelons *seuil de fonctionnements*. Formellement, le seuil de fonctionnements $\mathbf{F}_{\text{sup}}(L_h)$ est défini comme suit :

$$\mathbf{F}_{\text{sup}}(L_h) = \langle \mathcal{R}_h(\mathcal{J}_1), \mathcal{R}_h(\mathcal{J}_2), \dots, \mathcal{R}_h(\mathcal{J}_p) \rangle \quad (3.32)$$

3.4.3.5 Les capacités réalisables

Une '*capacité réalisable*' est un sous-ensemble de tous les vecteurs de fonctionnements possibles compatibles avec un ensemble donné de commodités. Formellement, à partir de chaque l ème vecteur de fonctionnements, noté $\mathbf{F}_l(L_h)$, associé au cluster L_h , nous construisons le sous-ensemble \mathbf{C}_h de l'ensemble de tous les vecteurs de fonctionnements possibles. \mathbf{C}_h représente la h ème capacité réalisable donné comme suit .

$$\mathbf{C}_h = \{ \mathbf{F}_l(L_h) \in \mathbb{C} : \mathbf{F}_l(L_h) \leq \mathbf{F}_{\text{sup}}(L_h) \} \quad (3.33)$$

où $h \in \{1, \dots, \mu\}$ et \mathbb{C} désigne l'ensemble non vide de tous les fonctionnements possibles. Remarquons que, $l \in \{1, \dots, l_h\}$ et l_h représente le nombre de vecteurs de fonctionnements du cluster L_h . Ainsi, nous introduisons la définition suivante.

Définition 3.4.3 *Un "vrai-pauvre" est un individu ou un ménage qui a un 'mauvais score' sur toutes ses commodités conflictuelles prise en compte simultanément pour évaluer ses 'capacité réalisable' dans la société dans laquelle il vit, dont la mesure de sa pauvreté est en conformité avec sa perception de sa situation actuelle.*

Notons qu'une capacité réalisable associée à un cluster donné reflète l'*aptitude* ou la *capacité* des ménages de ce cluster à accomplir un sous-ensemble de fonctionnements parmi les différentes alternatives de fonctionnements possibles.

3.4.3.6 Les capacités étendues

Certaines commodités (tels que le revenu) sont plutôt associées à l'ensemble des dimensions de bien-être. Nous appelons ces derniers '*commodités génériques*'.

Définition 3.4.4 *Une commodité générique est une commodité susceptible d'améliorer toute dimension du bien-être.*

Les capacités réalisables peuvent être étendues en utilisant les commodités génériques, lesquelles permettent d'augmenter certains niveaux de dimensions de bien-être de certains vecteurs de fonctionnements. Ainsi, les capacités réalisables sont transformés pour obtenir ce que nous avons appelé '*capacités étendues*'.

Formellement, soit $\xi_1^u > \xi_2^u > \dots > \xi_{t_u}^u$ les niveaux réalisables ordonnés pour tout $u \in \{1, \dots, p\}$. Soit $\mathbf{F}_{\text{sup}}(L_h)$ le *seuil de fonctionnements* de la capacité réalisable \mathbf{C}_h tels que nous ayons $\mathbf{F}_{\text{sup}}(L_h) = \langle \xi_{\mathbf{k}_1}^1, \xi_{\mathbf{k}_2}^2, \dots, \xi_{\mathbf{k}_p}^p \rangle$. De là, $\xi_{\mathbf{k}_u}^u \in \{\xi_1^u, \xi_2^u, \dots, \xi_{t_u}^u\}$ pour tout $\mathbf{k}_u \in \{1, \dots, t_u\}$ et $u \in \{1, \dots, p\}$. Nous posons $\xi_{\mathbf{k}_{\text{inf}}}^u = \min_{v \in \{1, \dots, t_u\}} \{\xi_v^u\}$ et $\xi_{\mathbf{k}_{\text{sup}}}^u = \max_{v \in \{1, \dots, t_u\}} \{\xi_v^u\}$. Une capacité étendue $\tilde{\mathbf{C}}_h$ est définie comme suite pour tous $h \in \{1, \dots, \mu\}$:

$$\text{Ext}(\mathbf{C}_h) = \tilde{\mathbf{C}}_h = \{\tilde{\mathbf{F}}_l(L_h) : l = 1, \dots, l_h\} \quad (3.34)$$

tel que

$$\tilde{\mathbf{F}}_l(L_h) = \mathbf{F}_{\text{sup}}(L_h) \cup \langle Q_{\mathbf{k}_1}^1, Q_{\mathbf{k}_2}^2, \dots, Q_{\mathbf{k}_p}^p \rangle \quad (3.35)$$

$$= \langle \xi_{\mathbf{k}_1}^1, \xi_{\mathbf{k}_2}^2, \dots, \xi_{\mathbf{k}_p}^p \rangle \cup \langle Q_{\mathbf{k}_1}^1, Q_{\mathbf{k}_2}^2, \dots, Q_{\mathbf{k}_p}^p \rangle \quad (3.36)$$

$$= \langle \Xi_{\mathbf{k}_1}^1, \Xi_{\mathbf{k}_2}^2, \dots, \Xi_{\mathbf{k}_p}^p \rangle \quad (3.37)$$

$$\text{où } \Xi_{\mathbf{k}_u}^u = \begin{cases} \xi_{\mathbf{k}_u}^u \cup Q_{\mathbf{k}_u}^u, & \text{si } Q_{\mathbf{k}_u}^u \neq \emptyset; \\ \xi_{\mathbf{k}_u}^u, & \text{sinon} \end{cases} \quad \text{pour tout } u \in \{1, \dots, p\}.$$

$$Q_{\mathbf{k}_u}^u = \{\xi_{\mathbf{k}_v}^v : \xi_{\mathbf{k}_v}^v > \xi_{\mathbf{k}_u}^u \text{ et } \mathbf{k}_v \in \{1, \dots, t_u\} \setminus \{\mathbf{k}_u\}\}.$$

En pratique, nous pouvons utiliser “*l’evidence-based policy*” (EBP) [voir [Davies, 2004](#), pour plus de détails] pour aborder cette question. Ceci implique l’utilisation de la collection de données sur la pauvreté ou le bien-être pour l’expérimentation, l’analyse quantitative et qualitative, l’utilisation des connaissances sur la pauvreté, la connaissance d’experts, les recherches nationales et internationales existantes, les statistiques existantes, les compétences des intervenants pour juger comment et dans quelle mesure les commodités génériques peuvent être utilisés en vue d’augmenter certains niveaux de bien-être des vecteurs de fonctionnement.

3.4.4 Exemple illustratif

3.4.4.1 Caractérisation des modalités

Considérant notre exemple illustratif précédent (Table 3.1) où $(\lambda_1, \lambda_2, \lambda_3) = (0.42, 0.25, 0.33)$ pour chaque cluster L_1, L_2 et L_3 , nous obtenons la Table (3.2) via le logiciel R¹ et la caractérisation modal-valuée $Y_j(L_h)$ avec $\xi(\beta_{jk})$ normalisée entre 0 et 1 par l’équation (3.20) : $Y_2(L_1) = \langle (N, 1.00); (Y, 0.00) \rangle$, $Y_3(L_1) = \langle (A, 0.06); (S, 0.85); (V, 0.00) \rangle$, $Y_4(L_1) = \langle (N, 1.00); (Y, 0.00) \rangle$, $Y_5(L_1) = \langle (A, 0.00); (F, 0.95); (V, 0.00) \rangle$, $Y_6(L_1) = \langle (A, 0.85); (N, 0.00); (S, 0.06) \rangle$, $Y_7(L_1) = \langle (Be, 0.00); (U, 0.00); (W, 1.00) \rangle$; $Y_2(L_2) = \langle (N, 0.00); (Y, 0.76) \rangle$, $Y_3(L_2) = \langle (A, 0.00); (S, 0.00); (V, 1.00) \rangle$, $Y_4(L_2) = \langle (N, 0.00); (Y, 0.76) \rangle$, $Y_5(L_2) = \langle (A, 0.00); (F, 0.00); (V, 1.00) \rangle$, $Y_6(L_2) = \langle (A, 0.00); (N, 1.00); (S, 0.00) \rangle$, $Y_7(L_2) = \langle (Be, 0.51);$

1. Téléchargement gratuit sur <http://www.r-project.org> [voir [R Development Core Team, 2011](#)]

$(U, 0.17); (W, 0.00)); Y_2(L_3) = \langle (N, 0.00); (Y, 0.89) \rangle, Y_3(L_3) = \langle (A, 1.00); (S, 0.00); (V, 0.00) \rangle, Y_4(L_3) = \langle (N, 0.00); (Y, 0.89) \rangle, Y_5(L_3) = \langle (A, 0.85); (F, 0.05); (V, 0.00) \rangle, Y_6(L_3) = \langle (A, 0.00); (N, 0.00); (S, 1.00) \rangle, Y_7(L_3) = \langle (Be, 0.37); (U, 0.48); (W, 0.00) \rangle$. Remarquons qu'il existe une distinction entre le critère modal-valoré $X_j(L_h)$ et la caractérisation modal-valorée $Y_j(L_h)$. Considérons par exemple $X_7(L_3)$ et $Y_7(L_3)$. $X_7(L_3)$ montre que : "pour la variable X_7 décrivant la 'situation économique' d'un ménage, 50% des ménages possède légalement la commodité 'Be' et la commodité 'U' dans le cluster L_2 ", tandis que $Y_7(L_3)$ montre que : "pour la variable X_7 décrivant la 'situation économique' d'un ménage, le degré de connexion de la commodité 'U' (égal à 0.48) est plus grand que le degré de connexion de la commodité 'Be' (égal à 0.37) dans le cluster L_3 ." Ainsi, la commodité 'U' (égal à 0.48) est plus connectée au cluster L_3 que la commodité 'Be' (égal à 0.37) ne l'est dans le même cluster L_3 . La notion de pertinence n'est pas mise en exergue par le critère modal-valoré $X_j(L_h)$ mais uniquement par la caractérisation modal-valorée $Y_j(L_h)$.

$F_{\mathcal{J}_u}$	Logement (\mathcal{J}_1)					Transport (\mathcal{J}_2)					Nutrition (\mathcal{J}_3)					
X	X_2		X_3			X_4		X_5			X_6			X_7		
β_{jk}	N	Y	A	S	V	N	Y	A	F	V	A	N	S	Be	U	W
$\rho(\beta_{jk})$	1.00	0.00	0.03	0.99	0.16	1.00	0.00	0.16	0.99	0.16	0.99	0.16	0.03	0.07	0.16	1.00
ζ_{jk}^1	1.00	0.00	0.20	1.00	0.00	1.00	0.00	0.00	0.83	0.00	1.00	0.00	0.20	0.00	0.00	1.00
π_{jk}^1	1.00	0.00	0.20	0.80	0.00	1.00	0.00	0.00	1.00	0.00	0.80	0.00	0.20	0.00	0.00	1.00
p_{jk}^1	0.42	0.58	0.42	0.33	0.25	0.42	0.58	0.25	0.50	0.25	0.33	0.25	0.42	0.33	0.25	0.42
$\xi(\beta_{jk})$	1.00	0.00	0.06	0.85	0.00	1.00	0.00	0.00	0.95	0.00	0.85	0.00	0.06	0.00	0.00	1.00
$\rho(\beta_{jk})$	0.16	0.84	0.16	0.26	1.00	0.16	0.84	0.38	0.09	1.00	0.26	1.00	0.16	0.76	0.38	0.16
ζ_{jk}^2	0.00	0.43	0.00	0.00	1.00	0.00	0.43	0.00	0.00	1.00	0.00	1.00	0.00	0.50	0.33	0.00
π_{jk}^2	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.67	0.33	0.00
p_{jk}^2	0.42	0.58	0.42	0.33	0.25	0.42	0.58	0.25	0.50	0.25	0.33	0.25	0.42	0.33	0.25	0.42
$\xi(\beta_{jk})$	0.00	0.76	0.00	0.00	1.00	0.00	0.76	0.00	0.00	1.00	0.00	1.00	0.00	0.51	0.17	0.00
$\rho(\beta_{jk})$	0.07	0.93	0.99	0.14	0.26	0.07	0.93	0.98	0.03	0.26	0.14	0.26	0.99	0.59	0.76	0.07
ζ_{jk}^3	0.00	0.57	0.80	0.00	0.00	0.00	0.57	1.00	0.17	0.00	0.00	0.00	0.80	0.50	0.67	0.00
π_{jk}^3	0.00	1.00	1.00	0.00	0.00	0.00	1.00	0.75	0.25	0.00	0.00	0.00	1.00	0.50	0.50	0.00
p_{jk}^3	0.42	0.58	0.42	0.33	0.25	0.42	0.58	0.25	0.50	0.25	0.33	0.25	0.42	0.33	0.25	0.42
$\xi(\beta_{jk})$	0.00	0.89	1.00	0.00	0.00	0.00	0.89	0.85	0.05	0.00	0.00	0.00	1.00	0.37	0.48	0.00
n_{jk}	5	7	5	4	3	5	7	3	6	3	4	3	5	4	3	5

TABLE 3.2 – Indices de la caractérisation des modalités

3.4.4.2 Les dimensions du bien-être

D'après l'exemple illustratif, nous considérons $F = \langle \text{Logement, Transport, Nutrition} \rangle$. Nous associons les sous-ensembles de commodités aux dimensions de bien-être. Par exemple, nous associons X_2 et X_3 à "Logement", X_4 et X_5 à "Transport", X_6 et X_7 à "Nutrition".

Remarque 3.4.3 Dans l'exemple illustratif l'unique commodité est le revenu donné par le critère X_1 .

3.4.4.3 Les niveaux de bien-être

Considérons notre exemple, nous pouvons associer les trois niveaux 'Bad' (Mauvais), 'Average' (Moyen) et 'Good' (Bien) à toutes dimensions du bien-être de sorte que $F_{\mathcal{J}_1}(L_h) = F_{\mathcal{J}_2}(L_h) = F_{\mathcal{J}_3}(L_h) = \{\mathbf{G}, \mathbf{A}, \mathbf{B}\}$, où $F_{\mathcal{J}_1} = \text{'Logement'}$, $F_{\mathcal{J}_2} = \text{'Transport'}$, $F_{\mathcal{J}_3} = \text{'Nutrition'}$.

3.4.4.4 Spécification des sous-ensembles S et U

Considérant l'exemple illustratif, le décideur peut partitionner l'ensemble des modalités de Satisfaction (et l'ensemble des modalités de Non satisfaction) des dimensions potentielles du bien-être \mathcal{J}_1 , \mathcal{J}_2 et \mathcal{J}_3 (appelée 'Logement', 'Transport' et 'Nutrition') comme suit :

$\mathbf{S}_{\mathcal{J}_1} = \{X_2.Y; X_3.V; X_3.A\}$, $\mathbf{S}_{\mathcal{J}_2} = \{X_4.Y; X_5.V; X_5.A\}$, $\mathbf{S}_{\mathcal{J}_3} = \{X_6.N; X_7.Be\}$,
 et $\mathbf{U}_{\mathcal{J}_1} = \{X_2.N; X_3.S\}$, $\mathbf{U}_{\mathcal{J}_2} = \{X_4.N; X_5.A; X_5.F\}$, $\mathbf{U}_{\mathcal{J}_3} = \{X_6.A; X_6.S; X_7.W\}$.
 La modalité Neutre est donnée par $\mathbf{N} = \{X_7.U\}$ mais pour des raisons empiriques nous l'avons rajouter à $\mathbf{S}_{\mathcal{J}_3}$ et $\mathbf{U}_{\mathcal{J}_3}$.

3.4.4.5 Construction des seuils de fonctionnements

Dans l'exemple (voir Table 3.4), nous considérons que ϑ_u est donnée par l'équation (4.9).

$$\mathbf{M}_h(\mathcal{J}_u) = \vartheta_u [\mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u}), \mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})] = \frac{\mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u}) - \mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})}{\max_{h=1, \dots, \mu} \{\mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u}), \mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})\}} \quad (3.38)$$

Le seuil $\mathbf{F}_{\text{sup}}(L_h)$ dans la Figure (3.3) représente le 'seuil de fonctionnements' que nous devons déterminer. En utilisant l'équation (4.2) et l'équation (4.3), nous calculons $\mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u})$ et $\mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})$ pour chaque cluster L_h , $h \in \{1, 2, 3\}$ et chaque dimension de bien-être $u \in \{1, 2, 3\}$ avec $w_j = 1$. Ceux-ci sont données dans la Table 3.3.

Clusters	Logement (\mathcal{J}_1)		Transport (\mathcal{J}_2)		Nutrition (\mathcal{J}_3)	
	$\mathbf{S}_{\mathcal{J}_1}$	$\mathbf{U}_{\mathcal{J}_1}$	$\mathbf{S}_{\mathcal{J}_2}$	$\mathbf{U}_{\mathcal{J}_2}$	$\mathbf{S}_{\mathcal{J}_3}$	$\mathbf{U}_{\mathcal{J}_3}$
\mathbf{C}_1	0.06	1.85	0.00	1.95	0.00	1.92
\mathbf{C}_2	1.76	0.00	1.76	0.00	1.68	0.17
\mathbf{C}_3	1.89	0.00	1.74	0.90	0.85	1.48

TABLE 3.3 – Évaluation de $\mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u})$ et $\mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})$

$$\mathcal{R}_h(\mathcal{J}_u) = \begin{cases} \text{Good (Bon),} & \mathbf{M}_h(\mathcal{J}_u) \geq \varepsilon_{1u} \\ \text{Average (Moyen),} & \mathbf{M}_h(\mathcal{J}_u) \geq \varepsilon_{2u} \\ \text{Bad (Mauvais),} & \mathbf{M}_h(\mathcal{J}_u) \geq \varepsilon_{3u} \end{cases} \quad (3.39)$$

$\mathbf{M}_h(\mathcal{J}_u)$ est donnée par l'équation (4.9) et $-\mathbf{M}_h(\varepsilon_u) \leq \varepsilon_{3u} < \varepsilon_{2u} < \varepsilon_{1u} \leq \mathbf{M}_h(\varepsilon_u)$ avec $\phi_h[\mathcal{J}_u] = \min \{ \mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u}); \mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u}) \}$; $\psi_h[\mathcal{J}_u] = \mathbf{C}_h(\mathbf{S}_{\mathcal{J}_u}) + \mathbf{C}_h(\mathbf{U}_{\mathcal{J}_u})$ et $\mathbf{M}_h(\varepsilon_u) = 1 - \frac{\phi_h[\mathcal{J}_u]}{\psi_h[\mathcal{J}_u]}$. Avec $\varepsilon_{1u} = 0.75$ et $\varepsilon_{2u} = 0.50$, nous calculons les mesures significantes donnée par la Table 3.4.

	\mathcal{J}_1	ε_1	\mathcal{J}_2	ε_2	\mathcal{J}_3	ε_3
\mathbf{M}_1	-0.95	0.97	-1.00	1.00	-1.00	1.00
\mathbf{M}_2	0.93	1.00	0.91	1.00	0.79	0.91
\mathbf{M}_3	1.00	1.00	0.43	0.66	-0.33	0.64

TABLE 3.4 – Les mesures significantes

Ainsi nous obtenons $\mathbf{F}_{\text{sup}}(L_1) = \langle \mathbf{B}, \mathbf{B}, \mathbf{B} \rangle$, $\mathbf{F}_{\text{sup}}(L_2) = \langle \mathbf{G}, \mathbf{G}, \mathbf{G} \rangle$ et $\mathbf{F}_{\text{sup}}(L_3) = \langle \mathbf{G}, \mathbf{B}, \mathbf{B} \rangle$.

3.4.4.6 Les vecteurs de fonctionnements

Compte tenu de notre exemple où nous avons associé les trois niveaux 'Bad', 'Average' et 'Good' à toutes les dimensions du bien-être telles que $F_{\mathcal{J}_1}(L_h) = F_{\mathcal{J}_2}(L_h) = F_{\mathcal{J}_3}(L_h) = \{ \mathbf{G}, \mathbf{A}, \mathbf{B} \}$ pour $F_{\mathcal{J}_1} = \text{'Logement'}$, $F_{\mathcal{J}_2} = \text{'Transport'}$, $F_{\mathcal{J}_3} = \text{'Nutrition'}$, le lème vecteur de fonctionnements associé au cluster h peut être donné par $\mathbf{F}_l(L_h) = \langle \mathbf{A}; \mathbf{B}; \mathbf{G} \rangle$ et un niveau d'accomplissement sur la dimension de bien-être 'Nutrition' est 'G'.

3.4.4.7 Le treillis de fonctionnements

La totalité de tous les vecteurs de fonctionnements possibles peut être visualisée par le treillis de fonctionnements tel que donné par la Figure (3.3) lequel montre

les vecteurs de fonctionnements dans un ordre du ‘meilleur’ (donné par $\langle G, G, G \rangle$) au pire (donné par $\langle B, B, B \rangle$). Si nous supposons par exemple que le symbole \rightarrow désigne la *relation de préférence* \succsim , $\langle G, A, G \rangle \rightarrow \langle G, B, A \rangle$ peut être interpréter par : “les conditions de vie offertes par $\langle G, A, G \rangle$ sont *au moins aussi préférables* que les conditions de vie offertes par $\langle G, B, A \rangle$ ”. La notation introduite dans la Figure (3.3) désigne une ‘*dominance naturelle*’ existante entre les fonctionnements. Cependant, quelques cas (tels que $\langle A, G, B \rangle$ vs $\langle G, B, A \rangle$) peuvent être difficile à trancher. Il revient au client d’en décider.

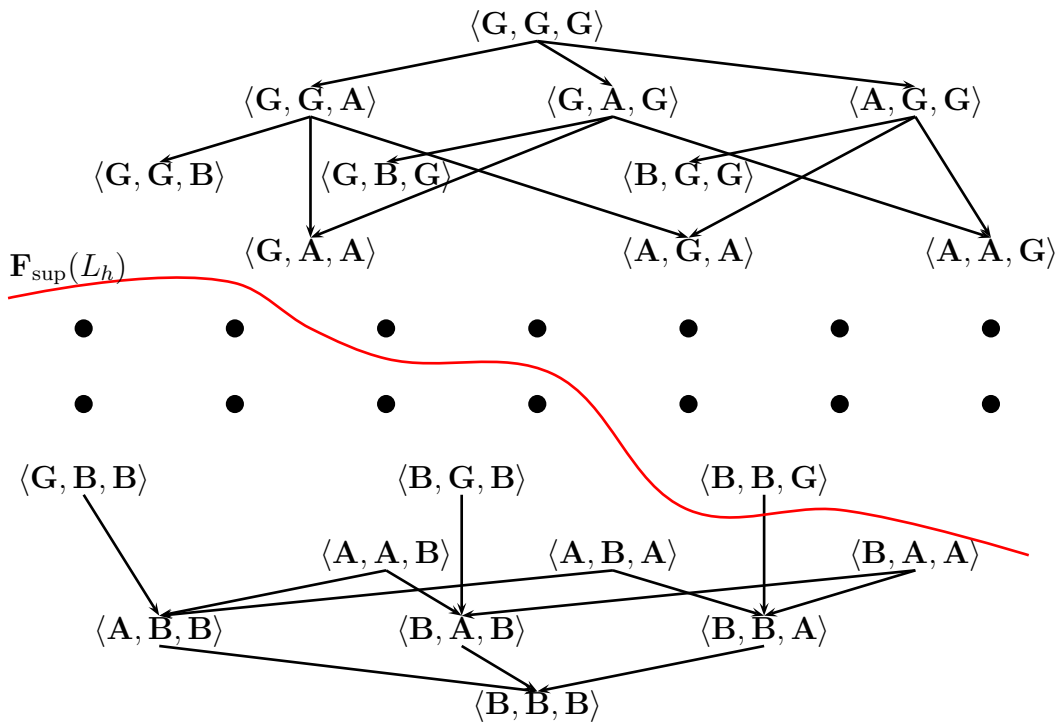


FIGURE 3.3 – Treillis de fonctionnements

3.4.4.8 Les capabilités réalisables

Les capabilités réalisables sont données dans la Table (3.5). Une analyse rapide des ménages de chaque groupe montre que la ‘pire’ capabilité réalisable C_1 ($F_{sup}(L_1) = \langle B, B, B \rangle$) est caractérisée par le fait que les chefs de ménages sont *en majorité* décrits par les modalités de Non-satisfaction U_{J_1} , U_{J_2} et U_{J_3} . Ce sont ceux que nous avons appelés les ‘vrais pauvres’ (voir définition 3.4.3). La ‘meilleure’ capabilité réalisable C_2 ($F_{sup}(L_2) = \langle G, G, G \rangle$) sont caractérisés par le fait que les chefs de ménages sont *en majorité* décrits par les modalités de Satisfaction S_{J_1} , S_{J_2} et S_{J_3} . La capabilité réalisable C_3 ($F_{sup}(L_3) = \langle G, B, B \rangle$) est caractérisée par les ménages qui sont en majorité décrits par les modalités de Satisfaction et de Non-satisfaction S_{J_1} , U_{J_2} and U_{J_3} . Ainsi, nous agrégeons la présence des com-

modités de Satisfaction ou de Non-satisfaction dans chaque cluster. Il est important de noter que, les mots 'pire' et 'meilleure' ne contiennent rien de numérique, mais uniquement la perception que le décideur attribue au bien-être et aux opportunités des ménages.

$$\begin{aligned}
 C_1 &= \{ \langle B, B, B \rangle \} \\
 C_2 &= \{ \langle G, G, G \rangle; \langle G, G, A \rangle; \langle G, A, G \rangle; \langle A, G, G \rangle; \dots \langle B, B, B \rangle \} \\
 C_3 &= \{ \langle G, B, B \rangle; \langle A, B, B \rangle; \langle B, B, B \rangle \}
 \end{aligned}$$

TABLE 3.5 – L'ensemble des capacités réalisables

3.4.4.9 Les capacités étendues

Les capacités réalisables sont transformées en capacités étendues en utilisant les commodités génériques (tels que le revenu sur l'exemple illustratif) lesquels permettent d'augmenter le(s) niveau(x) de bien-être de certains fonctionnements réalisables. Ceci soulève deux questions principales : Comment faire pour augmenter le(s) niveau(x) de bien-être de certains fonctionnements réalisables tout en évitant les contradictions et les incohérences qu'une telle opération peut susciter ? Comment traiter la question de légitimité et de rationalité d'une telle procédure ?

L'analyste doit déterminer les niveaux réalisables de chaque dimension du bien-être qu'il peut améliorer (ou augmenter) en utilisant les informations préférentielles fournies par le décideur. Considérant l'ensemble des capacités réalisables indiquées dans la Table (3.5), nous remarquons qu'il est possible d'améliorer les capacités réalisables C_3 sur la dimension de bien-être \mathcal{J}_2 représentant le 'Transport'. En fait, une observation rapide montre que $M_3(\mathcal{J}_3) < M_3(\mathcal{J}_2) < \varepsilon_v = 0.50$. De là, on peut être tenté de conclure que "les conditions de vie (ou le bien-être) offertes par C_1 et C_3 sur la dimension de bien-être \mathcal{J}_2 'est équivalent' aux conditions de vie (ou au bien-être) offertes par C_1 et C_3 sur la dimension de bien-être \mathcal{J}_3 ". La matrice modal-valuée donnée par l'équation (3.15) montre que $X_1(L_1) = [200, 1000]$ alors que $X_1(L_3) = [1500, 2800]$. Comme $M_3(\mathcal{J}_2) > M_3(\mathcal{J}_3)$ et $X_1(L_3) > X_1(L_1)$ (c'est-à-dire $M_3(\mathcal{J}_2) > M_3(\mathcal{J}_3)$ la Table (3.4) pour la commodité générique 'revenu' appartenant à l'intervalle $[1500, 2800]$), nous pouvons intuitivement augmenter le niveau réalisable de la dimension de bien-être \mathcal{J}_2 de L_3 de B à A. Pour la capacité réalisable C_3 , nous avons

$$F_l(L_3) = \langle \xi_{k_1}^1, \xi_{k_2}^2, \xi_{k_3}^3 \rangle \cup \langle Q_{k_1}^1, Q_{k_2}^2, Q_{k_3}^3 \rangle \quad (3.40)$$

$$= \langle \Xi_{k_1}^1, \Xi_{k_2}^2, \Xi_{k_3}^3 \rangle \quad (3.41)$$

où $\Xi_{k_1}^1 = \xi_{k_1}^1 = \mathbf{G}$ et $Q_{k_1}^1 = \emptyset$; $\Xi_{k_2}^2 = \xi_{k_2}^2 \cup Q_{k_2}^2 = \{\mathbf{B}\} \cup \{\mathbf{A}\} = \mathbf{A}$ et $Q_{k_1}^1 = \{\mathbf{A}\}$; $\Xi_{k_3}^3 = \xi_{k_3}^3 = \mathbf{B}$ et $Q_{k_3}^3 = \emptyset$. De là, la capabilité étendue de la capabilité réalisable C_3 est donnée dans la Table (3.6). Remarquons qu'une telle transformation des capabilités réalisables aux capabilités étendues dépend de l'information préférentielle fournie par le décideur et les faits scientifiques d'une expérience empirique.

$$\begin{aligned}\tilde{C}_1 &= \{\langle \mathbf{B}, \mathbf{B}, \mathbf{B} \rangle\} \\ \tilde{C}_2 &= \{\langle \mathbf{G}, \mathbf{G}, \mathbf{G} \rangle; \langle \mathbf{G}, \mathbf{G}, \mathbf{A} \rangle; \langle \mathbf{G}, \mathbf{A}, \mathbf{G} \rangle; \langle \mathbf{A}, \mathbf{G}, \mathbf{G} \rangle; \dots \langle \mathbf{B}, \mathbf{B}, \mathbf{B} \rangle\} \\ \tilde{C}_3 &= \{\langle \mathbf{G}, \mathbf{A}, \mathbf{B} \rangle; \langle \mathbf{G}, \mathbf{B}, \mathbf{B} \rangle; \langle \mathbf{A}, \mathbf{A}, \mathbf{B} \rangle; \langle \mathbf{A}, \mathbf{B}, \mathbf{B} \rangle; \langle \mathbf{B}, \mathbf{A}, \mathbf{B} \rangle; \langle \mathbf{B}, \mathbf{B}, \mathbf{B} \rangle\}\end{aligned}$$

TABLE 3.6 – L'ensemble des capabilités étendues

3.5 Implémentation des actions concrètes

Le lecteur doit noter que la construction d'une mesure signifiante doit toujours précéder l'élaboration des politiques et leur mise en œuvre, du fait que l'analyse des conséquences positives et négatives des politiques potentielles devrait être fait afin de fournir les informations appropriées au décideur. Ces informations sont cruciales pour les décideurs dans le sens où elles leur permettent de sélectionner des politiques adéquates pour chaque groupe de ménages. En d'autres termes, la raison d'être d'une mesure signifiante est liée à sa capacité d'aider à la conception et l'identification d'alternatives préférables à l'égard des questions politiques complexes. La mesure signifiante permet de faire des '*analyses*' parmi les alternatives. Les '*analyses*', dans notre contexte, impliquent mettre en lumière les valeurs et jugements implicites en adéquation avec les informations sur les préférences du décideur et l'identification des clusters qui sont en relation avec elles. Les clusters qui sont en contradiction avec les choix politiques alternatifs ne doivent pas être pris en considération.

Supposons que le décideur établisse par exemple un ensemble d'actions suivant : programme d'amélioration du logement (comme l'approvisionnement en infrastructures d'assainissement, l'approvisionnement en eau potable), programme d'amélioration de la nutrition, le soutien aux micro-crédits, etc. Considérant notre exemple illustratif, faire des '*analyses*' impliquent d'établir que la politique proposée par le décideur ne concerne que les ménages mal-logés et mal-nourris c'est-à-dire L_1 , et les ménages mal-nourris et acceptablement transportés c'est-à-dire L_2 . Remarquons que les ménages de L_2 ne sont soumis à aucun programme d'amélioration particulier. Ainsi, l'identification des clusters qui sont en accord avec la

politique établie par le décideur peut conduire à la mise en œuvre des politiques suivantes : les ménages de L_1 sont soumis à l'alternative (i), les ménages de L_3 sont soumis à l'alternative (ii) et les ménages de L_2 sont soumis à l'alternative (iii) comme indiquée ci-dessous :

- (i) Les ménages mal-logés et mal-nourris sont soumis au programme d'amélioration du logement (comme l'approvisionnement en infrastructures d'assainissement, l'approvisionnement en eau potable) et à un programme d'amélioration de la nutrition totalement pris en charge par le gouvernement
- (ii) Les ménages mal-nourris et acceptablement transportés sont soumis à un programme d'amélioration de la nutrition spécifique totalement prise en charge par le gouvernement si elles sont également mal-logés, mais ils doivent contribuer à hauteur de 70% des coûts s'ils sont acceptables logés ou bien-logés.
- (iii) Les ménages bien-logés, bien-transportés et bien-nourris ne sont soumis à aucune politique particulière, mais ils peuvent bénéficier d'une réduction d'impôt s'ils contribuent ou soutiennent financièrement le gouvernement dans l'application des actions concrètes établies dans (i) et (ii).

3.6 Discussion

L'approche par les capacités développée par Sen [1985] a fait l'objet de plusieurs critiques dans la littérature [see Navarro, 2000; Pogge, 2002; Nussbaum, 1987] dont les plus virulentes sont celles de Bénicourt [2004, 2006]. De manière générale, ces critiques évoquent l'ambiguïté de la théorie des capacités de Sen et des doutes quant à son opérationnalisation. La question qui suscite le débat est la suivante : *comment évaluer les capacités des individus ?* La nouvelle méthodologie développée dans cette thèse propose une réponse à cette question cruciale. Elle montre comment nous pouvons évaluer les capacités des individus de manière à prendre des décisions bien-informées au sujet des politiques, des programmes et des projets.

Ce chapitre 3 a consisté à la présentation de la première phase de notre méthodologie. En résumé, cette phase consiste à une classification non supervisée d'un échantillon de la population Ω dans le but de trouver une segmentation convenable et valide de la population en classes ayant des commodités socio-économiques homogènes. Pour atteindre cet objectif, nous utilisons une technique statistique de classification (le clustering) pour la découverte des regroupements naturels des ménages en faisant des comparaisons quantitatives de leurs commodités. Cette phase implique également la construction d'indices, le suivi-évaluation des capacités des ménages de chaque groupe et l'affectation des actions concrètes à chaque groupe identifié. Ainsi, les groupes avec les politiques correspondantes sont le résultat des préférences exprimées par les décideurs et reflètent la situation réelle des

populations concernées.

Les mesures multidimensionnelles significatives de la pauvreté (MDPMs) développées dans cette thèse sont un instrument d'élaboration et d'implémentation des politiques pour un suivi régulier et le diagnostic des problèmes sociaux. Cet instrument peut également être utilisé pour contrôler et évaluer les politiques d'inclusion sociale. Les mesures significatives qui en résultent sont *théoriquement solide* (du fait qu'elles respectent la nature des informations manipulées), *opérationnellement complète* (du fait qu'elles sont conçues pour répondre aux préoccupations de la population en termes de politiques) et produisent des résultats *légitimes* (car reflètent à la fois la réalité du terrain et les préférences du décideur). Cette méthodologie propose des mesures significatives satisfaisant trois compétences fondamentales : *compétence cognitive*, *compétence analytique*, *compétence prédiction des conséquences et évaluation des résultats*.

- COMPÉTENCE COGNITIVE : l'étape cognitive permet de comprendre quelle est la situation et d'identifier (à l'aide de la classification) les différents types de pauvreté au sein de la société à laquelle on peut plus tard associer des politiques adéquates. Du fait qu'il est plus logique d'avoir des différentes classes de personnes différemment pauvres, la table des données doit être organisée en classes de sorte que chaque classe se compose de ménages *similaires* entre eux en termes de distribution de commodités et *dissimilaires* aux ménages des autres classes. Il est plus vraisemblable de parler de politiques efficaces si elles sont bien ciblées. En guise d'exemple, la base de données ASSL 2007 du Burkina Faso (voir Annexe A) montre huit classes de ménages différemment pauvres données par $L = \{L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8\}$ mettant en exergue que la classe L_4 (de capacité réalisable C_4) est la plus déprivée c'est-à-dire la plus prioritaire en termes de politiques d'intervention ou d'implémentation des politiques.
- COMPÉTENCE ANALYTIQUE : l'étape analytique vise à identifier les commodités qui décrivent cette situation. Elle consiste à la description et la caractérisation des différentes classes qui ont été découvertes dans l'étape cognitive. Il s'agit alors dans cette étape d'identifier et de définir le critère modal-valué $X_j(L_h)$ et la caractérisation modal-valuée $Y_j(L_h)$ dans chaque cluster. Cela implique également une analyse des données avec pour but principal la mise en évidence des relations existantes entre les clusters, entre les commodités qui les caractérisent, entre les clusters et les variables. Ces relations, généralement difficiles à déceler dans de très grandes bases de données, sont parfois indispensables pour établir les recommandations finales. Par exemple, considérant notre base de données ASSL 2007 du Burkina Faso (voir Annexe A), les 'pires' capacités réalisables (soit $\{C_3; C_2; C_4\}$) sont caractérisées par le fait que le chef du ménage est *Agriculteur*, tandis que les 'meilleures' (soit $\{C_7; C_8; C_1\}$) sont caractérisées par le fait que le chef

de ménage est un *Employé du Secteur Public*. Cette information est crucial du fait qu'elle permet de "comprendre" chaque classe à partir des commodités des ménages, d'identifier les déprivations éventuelles, et de cibler les classes auxquelles nous pourrions associer les politiques adéquates (voir la représentation (A.1) dans Annexe A).

- **COMPÉTENCE PRÉDICTION DES CONSÉQUENCES ET ÉVALUATION DES RÉSULTATS** : du fait qu'une politique est considéré comme un ensemble d'actions (ou alternatives) que notre client souhaiterait entreprendre dans une région donnée, il est crucial d'explorer toutes les alternatives de chaque politique afin d'analyser les conséquences des différentes politiques possibles qui doivent être implémentées afin d'améliorer les conditions de vie des ménages. L'étape prédiction des conséquences et évaluation des résultats vise à associer un ensemble d'actions appropriées aux classes de manière à rendre possible et facile l'évaluation de l'efficacité des politiques actuelles et de déterminer si les conditions de vie des personnes s'améliorent ou pas (*'le suivi'*). Cela conduit à évaluer l'efficacité des diverses politiques possibles en mesurant si une politique particulière produira un impact positif sur le bien-être des personnes dans le future lorsqu'elle sera traduite en actions concrètes. Considérant, par exemple, la base de données ASSL 2007 du Burkina Faso (voir Annexe A) et en supposant que le décideur établit un ensemble d'actions, disons : programme d'amélioration du logement (comme l'approvisionnement en infrastructures d'assainissement, l'approvisionnement en eau potable, approvisionnement en électricité), programme d'éducation, programme d'amélioration de la nutrition, programme d'amélioration de la santé, programme d'amélioration du transport, le soutien au micro-crédit. Les classes de ménages $\{L_2, L_4\}$ (avec pour les capacités étendues $\{\tilde{C}_2, \tilde{C}_4\}$) peuvent bénéficier de tous les programmes d'amélioration des conditions de vie, tandis que la classe de ménage $\{L_3\}$ (avec pour capacités étendues $\{\tilde{C}_3\}$) n'a pas besoin de soutien au micro-crédit.

3.7 Quelques propriétés et axiomes

On peut montrer aisément que $\tilde{C}_h, h = 1, \dots, \mu$ (voir l'équation 3.33) satisfont aux propriétés et axiomes mathématiques développés dans Xu [2002]; Echávarri et Permanyer [2008]. Soit \mathbb{C} l'ensemble non vide de tous les vecteurs de fonctionnements possibles qui sont

Propriété 3.7.1 NON DÉGÉNÉRÉS : une capacité étendue $\tilde{C}_h \subset \mathbb{C}$ est dite non dégénérée si et seulement s'il existe $\mathbf{F}_l(L_h) = \langle F_{\mathcal{J}_1}(L_h); \dots; F_{\mathcal{J}_p}(L_h) \rangle_l \in \tilde{C}_h$ telle que $\mathcal{J}_u(L_h) > \varepsilon_u$ pour tout $u = 1, \dots, p$, où ε_u désigne le "plus petit fonctionnement" que nous l'appellerons le "zero-fonctionnement".

La propriété (3.7.1) indique que le décideur peut imposer une *limite inférieure* (“*le plus petit vecteur de fonctionnements*”) selon les politiques qu’il voudrait entreprendre dans une region donnée.

Propriété 3.7.2 COMPRÉHENSIFS : une capabilité étendue $\tilde{\mathbf{C}}_h \subset \mathbb{C}$ est dite *compréhensif* si et seulement si, pour tout $\mathbf{F}(L_{h_1}) = \langle F_{\mathcal{J}_1}(L_{h_1}); \dots; F_{\mathcal{J}_p}(L_{h_1}) \rangle$, $\mathbf{F}(L_{h_2}) = \langle F_{\mathcal{J}_1}(L_{h_2}); \dots; F_{\mathcal{J}_p}(L_{h_2}) \rangle \in \mathbb{C}$ avec $h_1, h_2 = 1, \dots, \mu$; si $F_{\mathcal{J}_u}(L_{h_1}) \geq F_{\mathcal{J}_u}(L_{h_2})$ pour tout $u = 1, \dots, p$, et $\mathbf{F}(L_{h_1}) \in \tilde{\mathbf{C}}_h$, alors $\mathbf{F}(L_{h_2}) \in \tilde{\mathbf{C}}_h$.

La propriété (3.7.2) indique que autant l’augmentation d’une dimension de bien-être dans un vecteur de fonctionnements quelconque est élevée, autant sa capabilité étendue est améliorée sur cette dimension.

Pour tout $\tilde{\mathbf{C}}_r, \tilde{\mathbf{C}}_s \subset \mathbb{C}$, nous disons que la capabilité étendue $\tilde{\mathbf{C}}_s$ est *entièrement contenue* dans $\tilde{\mathbf{C}}_r$ si pour tout $\mathbf{F}_b(L_s) \in \tilde{\mathbf{C}}_s$, il existe $\mathbf{F}_a(L_r) \in \tilde{\mathbf{C}}_r$ tel que $\mathbf{F}_a(L_r) > \mathbf{F}_b(L_s)$, avec $a \in \{1, \dots, l_r\}$ et $b \in \{1, \dots, l_s\}$. Ainsi, $\tilde{\mathbf{C}}_s$ est *entièrement contenue* dans $\tilde{\mathbf{C}}_r$ ssi $\mathbf{F}_{\text{sup}}(L_r) > \mathbf{F}_{\text{sup}}(L_s)$ et $\tilde{\mathbf{C}}_r \succ \tilde{\mathbf{C}}_s$.

La relation binaire \succsim définit sur \mathbb{C} satisfait la *réflexivité* c’est-à-dire pour tout $\tilde{\mathbf{C}}_r \subset \mathbb{C}$, $\tilde{\mathbf{C}}_r \succsim \tilde{\mathbf{C}}_r$ et la *transitivité* c’est-à-dire pour tout $\tilde{\mathbf{C}}_r, \tilde{\mathbf{C}}_s, \tilde{\mathbf{C}}_z \subset \mathbb{C}$, si $\tilde{\mathbf{C}}_r \succsim \tilde{\mathbf{C}}_s$ et $\tilde{\mathbf{C}}_s \succsim \tilde{\mathbf{C}}_z$ alors $\tilde{\mathbf{C}}_r \succsim \tilde{\mathbf{C}}_z$. Lorsque \succsim satisfait la *complétude*, c’est-à-dire pour tout $\tilde{\mathbf{C}}_r, \tilde{\mathbf{C}}_s \subset \mathbb{C}$ avec $\tilde{\mathbf{C}}_r \neq \tilde{\mathbf{C}}_s$, $\tilde{\mathbf{C}}_r \succsim \tilde{\mathbf{C}}_s$ ou $\tilde{\mathbf{C}}_s \succsim \tilde{\mathbf{C}}_r$, la *réflexivité* et la *transitivité*, alors \succsim est appelé un *ordre*.

La relation binaire \succsim sur \mathbb{C} satisfait les axiomes suivants :

Axiome 3.7.1 MONOTONICITÉ ssi : $\forall \tilde{\mathbf{C}}_r, \tilde{\mathbf{C}}_s \subset \mathbb{C}$, si $\tilde{\mathbf{C}}_s \subseteq \tilde{\mathbf{C}}_r$ alors $\tilde{\mathbf{C}}_r \succsim \tilde{\mathbf{C}}_s$, et si $\tilde{\mathbf{C}}_s \subset \tilde{\mathbf{C}}_r$, alors $\tilde{\mathbf{C}}_r \succ \tilde{\mathbf{C}}_s$.

La propriété de la monotonie indique que si la capabilité étendue $\tilde{\mathbf{C}}_s$ est un sous-ensemble de la capabilité étendue $\tilde{\mathbf{C}}_r$, alors les conditions de vie offertes par $\tilde{\mathbf{C}}_r$ sont au moins aussi préférables que les conditions de vie offertes par $\tilde{\mathbf{C}}_s$, et si $\tilde{\mathbf{C}}_s$ est *entièrement contenue* dans $\tilde{\mathbf{C}}_r$, alors $\tilde{\mathbf{C}}_r$ offre des conditions de vie au moins aussi plus élevées que ceux de $\tilde{\mathbf{C}}_s$. Autrement dit, les ménages appartenant au cluster L_r de capabilité étendue $\tilde{\mathbf{C}}_r$ possèdent des opportunités au moins aussi plus élevées que ceux du cluster L_s . Cette propriété met en évidence le fait qu’il est toujours désirable d’avoir une capabilité étendue plus large et reflète le degré de liberté des ménages d’un cluster à atteindre l’étendu des opportunités dont disposent ces ménages.

Preuve.

Soit $\tilde{\mathbf{C}}_r, \tilde{\mathbf{C}}_s \subset \mathbb{C}$ tel que $\tilde{\mathbf{C}}_s \subseteq \tilde{\mathbf{C}}_r$. Alors, pour tout $\mathbf{F}_b(L_s) \in \tilde{\mathbf{C}}_s$, il existe $\mathbf{F}_a(L_r) \in \tilde{\mathbf{C}}_r$ tel que $\mathbf{F}_a(L_r) \geq \mathbf{F}_b(L_s)$. Prenant par exemple $\mathbf{F}_a(L_r) = \mathbf{F}_{\text{sup}}(L_r)$, nous avons $\mathbf{F}_{\text{sup}}(L_r) \geq \mathbf{F}_b(L_s)$ pour tout $\mathbf{F}_b(L_s) \in \tilde{\mathbf{C}}_s$. Ainsi, $F_{\mathcal{J}_u}(L_r) \geq F_{\mathcal{J}_u}(L_s)$ pour tout $u \in \{1, \dots, p\}$, et par l’équation (3.29) nous obtenons $\sum_{u=1}^p \mathbf{M}_r(\mathcal{J}_u) \geq$

$\sum_{u=1}^p M_s(\mathcal{J}_u)$. Donc, $\tilde{C}_r \succ \tilde{C}_s$. En outre, $\tilde{C}_s \subset \tilde{C}_r$ implique \tilde{C}_s est entièrement contenue dans \tilde{C}_r c'est-à-dire $F_{\text{sup}}(L_r) > F_{\text{sup}}(L_s)$ et $\tilde{C}_r \succ \tilde{C}_s$ ■

Axiome 3.7.2 DÉsirABILITÉ DE CHAQUE FONCTIONNEMENT **ssi** : $\forall \tilde{C}_r \subset \mathbb{C}$ et pour tout $u = 1, \dots, p$, $\exists F_{\mathcal{J}_u}$ avec $\varepsilon_u < F_{\mathcal{J}_u} < \max_{\mathcal{J}_{u'}} \{F_{\mathcal{J}_{u'}} \in \mathbf{F}_l \text{ t.q. } \mathbf{F}_l \in \tilde{C}_r\}$ telle que $\tilde{C}_r \succ \tilde{C}_r \cap \{\mathbf{F}_z \in \mathbb{C} \text{ t.q. } F_{\mathcal{J}_v} \in \mathbf{F}_z \text{ et } F_{\mathcal{J}_v} \leq F_{\mathcal{J}_u}\}$ pour tout $v = 1, \dots, p$.

Cette axiome spécifie que, les conditions de vie offertes par une capabilité étendue diminuent lorsque nous retirons tous les vecteurs de fonctionnements ayant plus d'une certaine quantité de fonctionnement $F_{\mathcal{J}_v}$. Ceci signifie que chaque vecteur de fonctionnements d'une capabilité étendue est important dans l'évaluation des conditions de vie.

Preuve.

Soit $\tilde{C}_r \in \mathbb{C}$ et $F_{\text{sup}}(L_r) \in \mathbb{C}$ son seuil de fonctionnements. Nous posons $\mathbf{F}_a = \{F_{\mathcal{J}_u} \in \mathbf{F}_l : \varepsilon_u < F_{\mathcal{J}_u} < \max_{\mathcal{J}_{u'}} \{F_{\mathcal{J}_{u'}} \in \mathbf{F}_l \text{ t.q. } \mathbf{F}_l \in \tilde{C}_r\}, 1 \leq u \leq p\}$ et $\mathbf{C}_s = \{\mathbf{F}_z \in \mathbb{C} : F_{\mathcal{J}_v} \leq F_{\mathcal{J}_u} \text{ t.q. } F_{\mathcal{J}_v} \in \mathbf{F}_z, F_{\mathcal{J}_u} \in \mathbf{F}_a, \text{ and } 1 \leq v \leq p\}$. Alors, nous avons $F_{\text{sup}}(L_s) = \mathbf{F}_a < F_{\text{sup}}(L_r)$. Ceci implique que $\tilde{C}_r \cap \mathbf{C}_s \subset \tilde{C}_r$. Delà, $\tilde{C}_r \succ \tilde{C}_r \cap \mathbf{C}_s$ ■

Axiome 3.7.3 DOMINANCE **ssi** : pour tout entier positif N , tout $\tilde{C}_r, \mathbf{C}_1, \dots, \mathbf{C}_N \subset \mathbb{C}$ t.q. $\tilde{C}_s = \cup_{q=1}^N \mathbf{C}_q$ et $\tilde{C}_s \subset \mathbb{C}$, si $\tilde{C}_r \succ \mathbf{C}_1 \succ \mathbf{C}_2 \succ \dots \succeq \mathbf{C}_N$, alors $\tilde{C}_r \succ \tilde{C}_s$.

La dominance nécessite que lorsque la capabilité étendue \tilde{C}_r offre des conditions de vie au moins aussi élevées que tout \mathbf{C}_q (avec $q \in \{1, \dots, N\}$), et que les conditions de vie offertes par \mathbf{C}_q sont comparables, alors \tilde{C}_r offre des conditions de vie au moins aussi élevées que ceux de $\cup_{q=1}^N \mathbf{C}_q$. Cette axiome place une importance implicite sur le niveau de vie actuel du ménage lors du rangement de ses conditions de vie : autant l'aspect liberté des conditions de vie est concerné, nous pouvons avoir la situation dans laquelle mettant tous les \mathbf{C}_q ensemble augmenterait la liberté de choix offerte par chacun d'eux.

Preuve.

Soit $\tilde{C}_r, \mathbf{C}_1, \dots, \mathbf{C}_N \subset \mathbb{C}$ telle que $\tilde{C}_s = \cup_{q=1}^N \mathbf{C}_q$ et $\tilde{C}_s \subset \mathbb{C}$. De fait que $\tilde{C}_r \succ \mathbf{C}_1 \succ \mathbf{C}_2 \succ \dots \succeq \mathbf{C}_N$, le seuil de fonctionnements $F_{\text{sup}}(L_s) \in \tilde{C}_s$ tel que $F_{\text{sup}}(L_s) = F_{\text{sup}}(L_1)$ avec $F_{\text{sup}}(L_1) \in \mathbf{C}_1$. Comme $\tilde{C}_r \succ \mathbf{C}_1$, nous avons trivialement $\tilde{C}_r \succ \tilde{C}_s$ ■

Axiome 3.7.4 CONTINUITÉ **ssi** : pour tout $\tilde{C}_r \subset \mathbb{C}$ et une séquence de capabilités étendues $\{\tilde{C}_q\}_{q=1, \dots, +\infty}$ telle que $\tilde{C}_q \subset \mathbb{C}$ pour tout q et $\mathbf{C} = \lim_{q \rightarrow +\infty} \tilde{C}_q \subset \mathbb{C}$, si $\tilde{C}_q \succ \tilde{C}_r$ pour tout q , alors $\mathbf{C} \succ \tilde{C}_r$.

La continuité impose que le rangement des conditions de vie donné par la relation binaire \succsim ne doit pas inverser brusquement lorsqu'on tend vers l'infini. Ceci signifie que si toute capacité étendue \tilde{C}_q dans la séquence $\{\tilde{C}_q\}_{q=1,\dots,+\infty}$ offre des conditions de vie au moins aussi préférables que ceux de la capacité étendue \tilde{C}_r , alors la limite à l'infini de cette séquence devra aussi offrir des conditions de vie au moins aussi élevées que ceux de la capacité étendue \tilde{C}_r .

Preuve.

Considérons la capacité étendue $\tilde{C}_r \subset \mathbb{C}$ et une séquence $\{\tilde{C}_q\}_{q=1,\dots,+\infty}$, $\mathbf{C} = \lim_{q \rightarrow +\infty} \tilde{C}_q \succsim \tilde{C}_q$ par définition. De là, si $\tilde{C}_q \succsim \tilde{C}_r$ pour tout q , nous avons trivialement $\mathbf{C} \succsim \tilde{C}_r$ ■

Méthode de filtrage multi-attributs

4.1 Introduction

L'analyse des données se réfère à l'ensemble des méthodes et algorithmes qui se préoccupe de la description de données multidimensionnelles pour la prise de décision. Ce domaine (des statistiques [voir [Diday et al., 1982](#)]) peut être divisé soit en *analyse exploratoire* ou en *analyse confirmatoire* basée sur des modèles appropriés aux sources de données disponibles. L'élément fondamental de ces méthodes et algorithmes est le regroupement ou la classification des mesures en se basant (i) soit sur la qualité de l'ajustement (ou la cohérence) à un modèle postulé, ou (ii) soit sur l'organisation des données en des groupes sensibles révélés par l'analyse. Les méthodes de classification sont divisées en deux groupes : *la classification non supervisée* (clustering) et *la classification supervisée* (discriminant analysis). La classification non supervisée [[Jain et Dubes, 1988](#); [Jain et al., 1999](#); [Berkhin, 2002](#)] consiste à un regroupement des individus en clusters tels que tous les individus du même cluster soient *similaires* entre eux et *non similaires* aux individus des autres clusters. La classification supervisée [[Weiss et Kulikowski, 1991](#); [Mclachlan, 2004](#); [Kotsiantis, 2007](#)] est une technique d'affectation qui se base sur une collection d'individus appartenant aux clusters prédéfinis pour affecter de nouveaux individus.

Traditionnellement, les méthodes de classification se réfèrent aux problèmes dans lesquels l'information recherchée est contenue dans la table des évaluations (table de données) des individus sur différentes variables. Mais, la table de données brutes n'est généralement pas suffisante pour prendre des décisions ou recommandations appropriées à cause de la grande taille des données portant sur des variables mixtes. Par conséquent, notre objectif dans ce chapitre consiste à montrer comment affecter les nouveaux ménages à un ou plusieurs clusters prédéfinis en nous basant sur une règle d'affectation que nous avons dérivé du système de préférences relationnelles développé par [Roy \[1985\]](#).

Nous proposons ainsi dans ce chapitre une nouvelle méthode d'affectation floue basée sur les principes de concordance et de discordance dérivés des concepts bien connus de majorité et de veto [see [Vincke, 1992](#); [Perny et Roy, 1992](#); [Roy, 1996](#); [Perny, 1998](#); [Bouyssou, 1996](#); [Bouyssou et al., 2000](#); [Tsoukiàs et al., 2002](#)]. Cette méthode commence par la conception et la validation de cluster diagnostic, suivi

de l'affectation floue de nouveaux ménages aux clusters prédéfinis, lesquels sont en réalité des actions concrètes à implementer en termes de politiques, programmes ou projets.

Ce chapitre est organisé comme suite. Nous introduisons en Section 4.2 le concept de cluster diagnostic. En Section 4.3, nous présentons la nouvelle méthode d'affectation floue basée sur les principes de concordance et de discordance. Un exemple illustratif est développé en Section 4.4.

4.2 Construction des clusters diagnostics

4.2.1 Position du problème

Soit \mathbf{x}_i le i ème vecteur de commodités du ménage i . Abusivement, on note $\mathbf{U} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ l'ensemble des n ménages (la population totale) à affecter aux différents clusters L_h , avec $h \in \{1, \dots, \mu\}$, où $n \leq \mathbf{n}$ tels que $|\mathbf{U}| = \text{card}(\mathbf{U}) = \mathbf{n}$ et $i \in \{1, \dots, \mathbf{n}\}$. Notons que, les ensembles Ω et \mathbf{U} sont finis et non vides. Nous supposons sans nuire à la généralité que $\mathbf{U} \subset X$ tel que $X = X_1 \times \dots \times X_m$.

Soit $\mathbf{L} = \{L_1, \dots, L_\mu\}$ l'ensemble des clusters obtenus au Chapitre 3 de la classification non supervisée de l'échantillon Ω . On suppose sans nuire à la généralité que l'ensemble \mathbf{L} a été validé par le décideur. Notre objectif consiste alors à fournir une solution au problème suivant : étant donné un ménage $\mathbf{x}_i \in \mathbf{U}$ (avec $1 \leq i \leq \mathbf{n}$) évalué sur m variables et une liste de clusters prédéfinis $\{L_1, \dots, L_\mu\}$ de capacités étendues $\{\tilde{\mathbf{C}}_1, \dots, \tilde{\mathbf{C}}_\mu\}$ caractérisée par une liste de clusters diagnostics $\{\mathbf{D}_1, \dots, \mathbf{D}_\mu\}$ respectivement.

$$\begin{aligned}
 L_1 &= \{x_{11}, \dots, x_{1n_1}\} \equiv \mathbf{D}_1 \quad \text{related to } \mathbf{C}_1 \\
 L_2 &= \{x_{21}, \dots, x_{2n_2}\} \equiv \mathbf{D}_2 \quad \text{related to } \mathbf{C}_2 \\
 &\vdots \\
 L_\mu &= \{x_{\mu 1}, \dots, x_{\mu n_\mu}\} \equiv \mathbf{D}_\mu \quad \text{related to } \mathbf{C}_\mu
 \end{aligned} \tag{4.1}$$

où x_{hk} est le k -ème ménage du h -ème cluster et n_h est la taille du h -ème cluster. Nous souhaitons calculer le degré d'appartenance de chaque ménage de \mathbf{U} à chaque cluster L_h , $h \in \{1, \dots, \mu\}$. Comment pouvons-nous traiter ce problème ?

4.2.2 Les clusters diagnostics

Les indices de pertinence, tels que définis par l'équation (3.18), sont des représentations numériques qui permettent d'effectuer des comparaisons entre les modalités. Cependant, ces indices peuvent être utilisés de façon plus appropriée pour capturer la "sensation" des ménages par rapport aux différents clusters. Ceci est possible grâce aux clusters diagnostics. Le cluster diagnostic est le résultat de

la transformation de la table des données de grande taille en un tableau récapitulatif dans le but d'acquérir des connaissances cruciales sur chaque cluster. C'est un outil visant à identifier et à décrire la nature et les raisons des différents types de pauvreté observés au sein de la population Ω afin de produire le diagnostic de la situation. Comme nous l'avons déjà dit dans le chapitre précédent, les individus sont plus ou moins pauvres et de différentes façons, il est donc plus logique d'élaborer un diagnostic de chaque cluster à partir duquel nous pouvons affecter de nouveaux ménages. Ainsi, le diagnostic des clusters consiste en une analyse approfondie des différents clusters de manière à identifier et à connaître les relations entre les commodités et les clusters. Chaque cluster issue d'un tel processus s'appelle "*cluster diagnostic*".

Formellement, soit \mathbf{d}_{hj} un sous-ensemble de \mathbf{X}^j ($\mathbf{d}_{hj} \subseteq \mathbf{X}^j$) avec $h \in \{1, \dots, \mu\}$ et $j \in \{1, \dots, m\}$. $\mathbf{D}_h \subseteq \mathbf{X}$ désigne le diagnostic du cluster L_h et \mathbf{d}_{hj} désigne le '*diagnostic*' du cluster L_h sur la j ème variable X_j . \mathbf{D}_h est abusivement appelé '*cluster diagnostic de L_h* ' et \mathbf{d}_{hj} '*diagnostic de L_h sur X_j* '. Ainsi, \mathbf{d}_{hj} est le *diagnostic* de la caractérisation modal-valuée $Y_j(L_h)$ et \mathbf{D}_h est le diagnostic du cluster modal-valué \mathbf{X}_h . Chaque diagnostic \mathbf{D}_h fournit une description des ménages du cluster L_h dotés des capacités étendues $\tilde{\mathbf{C}}_h$. L'idée consiste alors à diviser la totalité des modalités en trois sous-ensembles : le sous-ensemble des modalités **Admis** $\mathbf{d}_{hj}^{\mathbf{A}}$, le sous-ensemble des modalités **Eligible** $\mathbf{d}_{hj}^{\mathbf{E}}$ et le sous-ensembles des modalités **Rejeté** $\mathbf{d}_{hj}^{\mathbf{R}}$. De là, un diagnostic \mathbf{d}_{hj} est défini comme suit :

$$\mathbf{d}_{hj} = \mathbf{d}_{hj}^{\mathbf{A}} \cup \mathbf{d}_{hj}^{\mathbf{E}} \cup \mathbf{d}_{hj}^{\mathbf{R}} = \{(\beta_{jk}, \chi_{jk}) : k = 1, \dots, s_j\} \quad (4.2)$$

et

$$\chi_{jk} = \begin{cases} \mathbf{Admis} & \text{if : } \xi(\beta_{jk}) \geq \tau_1 \\ \mathbf{Eligible} & \text{if : } \tau_0 \leq \xi(\beta_{jk}) < \tau_1 \\ \mathbf{Rejeté} & \text{if : } 0 \leq \xi(\beta_{jk}) < \tau_0 \end{cases} \quad \text{with } 0 \leq \tau_0 < \tau_1 \leq 1 \quad (4.3)$$

où s_j est le nombre de valeurs actuellement prise par le domaine D_j . Le diagnostic \mathbf{D}_h , donné par l'équation (4.4), est l'ensemble des diagnostics \mathbf{d}_{hj} pour tout $j \in \{1, \dots, m\}$.

$$\mathbf{D}_h = \langle \mathbf{d}_{h1}, \mathbf{d}_{h2}, \dots, \mathbf{d}_{hm} \rangle \quad (4.4)$$

Remarque 4.2.1 *Le sous-ensemble des modalités **Admis** est formé des modalités qui caractérisent 'fortement' un cluster donné, tandis que le sous-ensemble des modalités **Eligible** est formé de ceux qui le caractérisent 'partiellement' et le sous-ensemble des modalités **Rejeté** est formé de ceux qui le caractérisent 'faiblement'.*

Les dépendances inhérentes à la structure des clusters et des données peuvent être parfaitement élucidées grâce un examen minutieux des clusters diagnostics \mathbf{D}_h

pour tout $h \in \{1, \dots, \mu\}$. Par ‘dépendances’, nous faisons référence aux dépendances logiques et hiérarchiques et non à la dépendance statistique qui peut ou ne pas exister entre les variables aléatoires X_j ou les clusters L_h .

4.3 La méthode d’affectation floue

Roy [1968] fût le premier à utiliser le principe de concordance et de discordance dans les méthodes d’aide à la décision. La famille des méthodes ELECTRE, bien connue dans la littérature, représente ses premiers travaux sur la méthodologie multicritère d’aide à la décision. Les principes de concordance et de discordance sont dérivés des concepts bien connus de la majorité et du droit de veto [voir Vincke, 1992; Perny et Roy, 1992; Roy, 1996; Perny, 1998; Bouyssou, 1996; Bouyssou *et al.*, 2000; Tsoukiàs *et al.*, 2002]. Nous traiterons dans cette section le problème d’affectation des ménages aux différents clusters prédéfinis en se basant sur le principe de concordance et de discordance. Dans la suite, nous allons souvent remplacer la terminologie ‘cluster’ par ‘capabilité étendue’ de manière à garder en esprit le fait qu’une capabilité étendue est plus “signifiante” qu’un cluster.

4.3.1 Principe concordance

Le principe de concordance [voir Roy, 1978; Perny, 1998] utilise la règle de majorité pour décider entre chaque pair \mathbf{x}_i et \mathbf{D}_h . Il représente l’existence d’une coalition “*significative*” de modalités pour laquelle “ \mathbf{x}_i appartient à la capabilité étendue $\tilde{\mathbf{C}}_h$ ” noté par $\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$.

Définition 4.3.1 Une relation \mathcal{R} est une sous-ensemble du produit cartésien $X \times \cup_{h=1}^{\mu} \mathbf{D}_h$ reliant un vecteur de commodités spécifique \mathbf{x}_i (avec $\mathbf{x}_i \in \mathbf{U}$) à un cluster diagnostic approprié \mathbf{D}_h .

Pour tout $\mathbf{x}_i \in \mathbf{U}$ et $h \in \{1, \dots, \mu\}$, cette relation notée $\mathcal{R}(\mathbf{x}_i, \mathbf{D}_h) \in \{0; 1\}$ prend la valeur booléenne 1 si une relation donnée est vrai et 0 sinon. En règle générale, la relation \mathcal{R} prend l’une des formes, \leq , \geq , $<$, $>$, \subset , \subseteq , \neq , $=$, \in , \notin . Ainsi, la relation $\mathcal{R} = \prod_{j=1}^m \mathcal{R}_j$ est définie comme étant le produit de (tout) nombre booléen de relations binaires simples reliant l’observation spécifique de X_j (X respectivement), $j \in \{1, \dots, m\}$, au diagnostic \mathbf{d}_{hj} (cluster diagnostic \mathbf{D}_h respectivement), $h = 1, \dots, \mu$. Elle est donnée par l’équation (4.5) comme suit :

$$\mathcal{R}(\mathbf{x}_i, \mathbf{D}_h) = \bigwedge_{1 \leq j \leq m} \mathcal{R}_j(x_i^j, \beta_{jk}) \quad (4.5)$$

où \bigwedge indique la conjonction logique ‘et’. De là, nous construisons deux indices qui seront utilisés pour calculer un indice global $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ mesurant l’adéquation globale avec la proposition “ \mathbf{x}_i appartient à la capabilité étendue $\tilde{\mathbf{C}}_h$ ” notée $\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h$:

Définition 4.3.2 L'indice d'admissibilité $\mathbf{A}(\mathbf{x}_i, \mathbf{D}_h)$ reflète le degré auquel “ \mathbf{x}_i appartient fortement à la capacité étendue $\tilde{\mathbf{C}}_h$ ” est en accord avec la réalité représentée par les évaluations de \mathbf{x}_i sur tous les attributs ‘valides’ et ses capacités étendues $\tilde{\mathbf{C}}_h$. Cet indice est donné par l'équation (4.6) comme suit :

$$\mathbf{A}(\mathbf{x}_i, \mathbf{D}_h) = \frac{\sum_{j \in \mathbf{T}[\mathbf{A}]} w_j \xi(\beta_{jk})}{\sum_{j \in \mathbf{T}[\mathcal{R}]} w_j \xi(\beta_{jk})} \quad (4.6)$$

où $\mathbf{T}[\mathbf{A}]$ et $\mathbf{T}[\mathcal{R}]$ sont donnés par l'équation (4.7) et l'équation (4.12) respectivement ; k et τ_1 sont données par les équations (4.2) et (4.3).

$$\mathbf{T}[\mathbf{A}] = \{j \in \mathbf{T}[\mathcal{R}] \ :: \ \langle \mathcal{R}_j(x_i^j, \beta_{jk}) = 1 \rangle \wedge \langle \xi(\beta_{jk}) \geq \tau_1 \rangle\} \quad (4.7)$$

Définition 4.3.3 L'indice d'éligibilité $\mathbf{E}(\mathbf{x}_i, \mathbf{D}_h)$ reflète le degré auquel “ \mathbf{x}_i appartient partiellement à la capacité étendue $\tilde{\mathbf{C}}_h$ ” est en accord avec la réalité représentée par les évaluations de \mathbf{x}_i sur tous les attributs ‘valides’ et ses capacités étendues $\tilde{\mathbf{C}}_h$. Cet indice est donné par l'équation (4.10) comme suit :

$$\mathbf{E}(\mathbf{x}_i, \mathbf{D}_h) = \frac{\sum_{j \in \mathbf{T}[\mathbf{E}]} w_j \xi(\beta_{jk})}{\sum_{j \in \mathbf{T}[\mathcal{R}]} w_j \xi(\beta_{jk})} \quad (4.8)$$

où $\mathbf{T}[\mathbf{E}]$ et $\mathbf{T}[\mathcal{R}]$ sont donné par les équations (4.9) et (4.12) respectivement.

$$\mathbf{T}[\mathbf{E}] = \{j \in \mathbf{T}[\mathcal{R}] \ :: \ \langle \mathcal{R}_j(x_i^j, \beta_{jk}) = 1 \rangle \wedge \langle \tau_0 \leq \xi(\beta_{jk}) < \tau_1 \rangle\} \quad (4.9)$$

Définition 4.3.4 L'indice de rejet $\mathbf{R}(\mathbf{x}_i, \mathbf{D}_h)$ reflète le degré auquel “ \mathbf{x}_i appartient faiblement à la capacité étendue $\tilde{\mathbf{C}}_h$ ” s'accorde avec la réalité représentée par les évaluations de \mathbf{x}_i sur tous les attributs ‘valides’ et ses capacités étendues $\tilde{\mathbf{C}}_h$. Cet indice est donné par l'équation (4.10) comme suit :

$$\mathbf{R}(\mathbf{x}_i, \mathbf{D}_h) = \frac{\sum_{j \in \mathbf{T}[\mathbf{R}]} w_j \xi(\beta_{jk})}{\sum_{j \in \mathbf{T}[\mathcal{R}]} w_j \xi(\beta_{jk})} \quad (4.10)$$

où $\mathbf{T}[\mathbf{R}]$ et $\mathbf{T}[\mathcal{R}]$ sont donnés par les équations (4.11) et (4.12) respectivement ; k , τ_0 et τ_1 sont donné par les équations (4.2) et (4.3).

$$\mathbf{T}[\mathbf{R}] = \{j \in \mathbf{T}[\mathcal{R}] \ :: \ \langle \mathcal{R}_j(x_i^j, \beta_{jk}) = 1 \rangle \wedge \langle 0 \leq \xi(\beta_{jk}) < \tau_0 \rangle\} \quad (4.11)$$

$$\mathbf{T}[\mathcal{R}] = \{j \in \mathcal{J} \ :: \ \langle \mathcal{R}_j(x_i^j, \beta_{jk}) = 1 \rangle \wedge \langle 0 \leq \xi(\beta_{jk}) \leq 1 \rangle\} \quad (4.12)$$

Définition 4.3.5 La coalition de concordance pour $\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h$ est le sous-ensemble flou $F_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h}$ pour tout $h = 1, \dots, \mu$ et pour tout ménage $\mathbf{x}_i \in \mathbf{U}$:

$$F_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h} = \{(\beta_{jk}, \mu_{jk}^h) : k = 1, \dots, s_j \text{ and } j \in \mathbf{T}[\mathcal{R}]\} \quad (4.13)$$

où s_j est le nombre de valeurs du domaine D_j et μ_{jk}^h est le degré d'appartenance de la modalité β_{jk} dans l'ensemble flou $F_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h}$.

Définition 4.3.6 L'indice global de concordance $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ est une relation binaire valuée entre le vecteur des commodités \mathbf{x}_i et la capacité étendue $\tilde{\mathbf{C}}_h$ mesurant l'importance globale de la coalition $F_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h}$ en utilisant la formule donnée par l'équation (4.14) comme suit :

$$\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) = \frac{\sum_{j \in \mathbf{T}[\mathbf{A}]} w_j \xi(\beta_{jk}^h) + \sum_{j \in \mathbf{T}[\mathbf{E}]} w_j \alpha_j \xi(\beta_{jk})}{\sum_{j \in \mathbf{T}[\mathcal{R}]} w_j \xi(\beta_{jk})} \quad (4.14)$$

avec $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) \geq \lambda$; $\{\alpha_j :: j \in \mathbf{T}[\mathbf{E}] \text{ et } 0 \leq \alpha_j \leq 1\}$ indique une système de pondération nécessaire à l'analyse de la sensibilité et $0.5 \leq \lambda \leq 1 - \frac{\min_{j \in \mathbf{T}[\mathcal{R}]} \xi(\beta_{jk})}{\sum_{j \in \mathbf{T}[\mathcal{R}]} \xi(\beta_{jk})}$.

Ainsi, l'indice de concordance est une fonction $H(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ satisfaisant trois exigences [voir Roy et Bouyssou, 1993] pour tout $\mathbf{x}_i \in \mathbf{U}$ et $h \in \{1, \dots, \mu\}$:

1. Le i ème ménage n'intervient dans le calcul $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ que par les seuls vecteur de commodités \mathbf{x}_i .
2. $0 \leq \mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) \leq 1$; nous avons :
 - $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) = 0$ lorsque la proposition “ \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ” n'est appuyée par aucun argument significatif susceptible de la valider.
 - $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) = 1$ lorsque la proposition “ \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ” est appuyée par des argument tout à fait convaincants permettant de ne pas douter de sa validité.
3. $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) > \mathbf{H}(\mathbf{x}_k, \tilde{\mathbf{C}}_h)$ lorsque la proposition “ \mathbf{x}_k appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ” est appuyée par des arguments qui la rendent plus crédible que la proposition “ \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ”.

Cet indice permet de construire un ‘continuum de niveau’ et d'évaluer combien la proposition “ \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ” est significative pour valider l'affectation. Notons que, la définition d'affectation ‘forte’ or appartenance ‘forte’ à un cluster donné L_h implique conjointement un degré d'admissibilité élevé (suffisamment grand) et un degré d'éligibilité suffisamment petit.

4.3.2 Principe de discordance

Le principe de discordance [voir Roy, 1978; Perny, 1998] indique de prendre en compte les minorités fortement contradictoires pour décider entre chaque paire $(\mathbf{x}_i, \mathbf{D}_h)$. Cet indice représente l'existence d'une "opposition significative" contre la proposition " \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ". Le rôle du test de discordance est de vérifier si certaines modalités s'opposent fortement à la proposition " \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ".

Définition 4.3.7 L'ensemble flou de discordance de $\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h$ est le sous-ensemble $D_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h}$ pour tout $h \in \{1, \dots, \mu\}$ et pour tout ménage $\mathbf{x}_i \in \mathbf{U}$:

$$D_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h} = \{(\mathbf{x}_i, \nu_i^h) : \mathbf{x}_i \in \mathbf{U}\} \quad (4.15)$$

où ν_i^h est une relation binaire valuée entre le vecteur des commodités \mathbf{x}_i et la capacité étendue $\tilde{\mathbf{C}}_h$ reflétant le degré auquel le i ème ménage est discordant dans le sous-ensemble flou $D_{\mathbf{x}_i \mathbf{H} \tilde{\mathbf{C}}_h}$. ν_i^h est donnée par :

$$\nu_i^h = 1 - \min \left(1, \max \left(0, \frac{A}{v_1^h} \right) \right) \times \min \left(1, \max \left(0, \frac{B}{v_2^h} \right) \right) \quad (4.16)$$

où $A = \mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) - \mathbf{R}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ et $B = \mathbf{A}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) - \mathbf{E}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) - \mathbf{R}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$; $v_1^h > 0$ ($v_2^h > 0$, respectivement) est le seuil de veto dans l'indice de concordance (l'indice d'admissibilité, respectivement). Ce seuil permet de limiter les compensations possibles entre les "avantages" et "inconvénients"

Remarque 4.3.1 Une simple solution inspirée de la méthode ELECTRE III [voir Roy, 1978] est donnée par l'indice de discordance $\mathbf{D}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) = \nu_i^h$ où $\mathbf{D}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) > \delta$ et $0.5 \leq \delta \leq 1$.

Notons que, l'indice de discordance est défini de telle manière que la présence d'au moins une contradiction forte (discordance forte) (telle que $\nu_i^h = 1$) est suffisante pour rendre l'indice de discordance positif. Les seuils λ et δ sont des seuils de concordance et de discordance globale. Notons que, l'affectation des seuils de veto résulte de la discussion entre le décideur et l'analyste.

L'indice de concordance (voir l'équation (4.14)) et l'indice de discordance (voir l'équation (4.16)) sont alors utilisés pour formuler la relation de préférence globale de type " \mathbf{x}_i appartient à la capacité étendue $\tilde{\mathbf{C}}_h$ ", noté par $\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$, telle que donné par l'équation (4.17). La proposition $\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ est valide si et seulement si la coalition des attributs en accord avec cette proposition est assez forte et il n'existe pas de coalition significative discordante avec elle.

$$\mathbf{x}_i \in \tilde{\mathbf{C}}_h \iff \mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) \wedge \neg \mathbf{D}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) \quad (4.17)$$

où les symboles \wedge et \neg sont respectivement la conjonction et la négation. L'opérateur \wedge peut être interprété en utilisant une t -norme T et l'opérateur \neg peut être interprété en utilisant la négation strict N , c'est-à-dire $N(x) = \phi^{-1}(1 - \phi)$ où ϕ est un automorphisme sur l'intervalle $[0; 1]$. L'équation (4.17) devient alors :

$$\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) = T(\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h), N(\mathbf{D}(\mathbf{x}_i, \tilde{\mathbf{C}}_h))) \quad (4.18)$$

Ceci conduit à construire le degré d'appartenance $\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) \in [0, 1]$ comme une fonction non-décroissante de $\mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$ et une fonction non-croissance de $\mathbf{D}(\mathbf{x}_i, \tilde{\mathbf{C}}_h)$. En guise d'illustration, nous avons choisi la solution proposée par Perny [1998] :

$$\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) = \min \{ \mathbf{H}(\mathbf{x}_i, \tilde{\mathbf{C}}_h), 1 - \mathbf{D}(\mathbf{x}_i, \tilde{\mathbf{C}}_h) \} \quad (4.19)$$

4.3.3 Règle d'affectation

Soit $\{\tilde{\mathbf{C}}_1, \dots, \tilde{\mathbf{C}}_\mu\}$ l'ensemble des capacités étendues. Chaque capacité étendue est caractérisée par un ensemble de diagnostic $\mathbf{D}_h \subseteq X_h$ contenant les '*modalités références*' qui décrivent les ménages d'un cluster L_h , $h = 1, \dots, \mu$. Dans cette thèse, nous avons supposé que les capacités étendues ne sont pas nécessairement ordonnées et ne forment pas nécessairement une partition des situations possibles. Ainsi, un 'nouveau' ménage peut être affecté à plusieurs clusters avec des degrés d'appartenance différents selon le cluster lorsque ses commodités coïncide avec plus d'un cluster diagnostic.

La règle d'affectation vise à affecter ou associer un ménage à sa capacité étendue appropriée en se basant sur une condition prédéfinie que nous pouvons exprimer comme suit : "*un ménage $\mathbf{x}_i \in \mathbf{U}$ est affecté à la capacité étendue $\tilde{\mathbf{C}}_h$, $h \in \{1, \dots, \mu\}$, si et seulement si le vecteur de commodités \mathbf{x}_i coïncide avec le cluster diagnostic \mathbf{D}_h de L_h c'est-à-dire il existe une relation de voisinage entre \mathbf{x}_i et les modalités décrivant les ménages du cluster L_h* ". Donc, la règle d'affectation peut être représenté par l'équation (4.20) comme suite :

$$\mathbf{x}_i \in \tilde{\mathbf{C}}_h \iff \mathbf{x}_i \in \vartheta(\mathbf{D}_h) \quad (4.20)$$

où $\vartheta(\mathbf{D}_h) = \{\mathbf{x}_i \in \mathbf{U} :: \mu_{\vartheta(\mathbf{D}_h)}(\mathbf{x}_i) \geq \alpha\}$ et α est un seuil proposé par le décideur. Pour tous ménages \mathbf{x}_i , $\mu_{\vartheta(\mathbf{D}_h)}(\mathbf{x}_i)$ représente le degré d'appartenance de \mathbf{x}_i au cluster L_h au vue de son cluster diagnostic \mathbf{D}_h .

Ainsi, pour toute relation d'appartenance évaluée ou non-évaluée \mathbf{B} utilisant le principe de concordance et de non-discordance, la règle d'affectation est donnée par l'équation suivante :

$$\mu_{\vartheta(\mathbf{D}_h)}(\mathbf{x}_i) = \vee(\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_1), \dots, \mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_\mu)) \quad (4.21)$$

où \vee est une co-norme. Nous avons adopté dans cette thèse la max co-norme. Ainsi, l'affectation du i ème ménage à la capacité étendue $\tilde{\mathbf{C}}_h$ (avec $h \in \{1, \dots, \mu\}$) est

déterminée par le degré d'appartenance du i th ménage de vecteur de commodités \mathbf{x}_i à la capacité étendue $\tilde{\mathbf{C}}_h$ lequel est mesuré par son plus proche voisin dans $\tilde{\mathbf{C}}_h$ (avec $h \in \{1, \dots, \mu\}$).

$$\mu_{\vartheta(\mathbf{D}_h)}(\mathbf{x}_i) = \max(\mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_1), \dots, \mathbf{B}(\mathbf{x}_i, \tilde{\mathbf{C}}_\mu)) \quad (4.22)$$

4.4 Exemple illustratif

Considérons notre exemple illustration donné par la Table 3.1 et supposons que les sept variables ont approximativement la même importance. De là, nous attribuons les seuils veto égale à 0.95 sur chaque indice de concordance, les seuils veto égale 0.40 sur chaque indice d'admissibilité pour tous les clusters. Les poids utilisés dans l'agrégation sont égaux pour tous les clusters bien que, dans d'autre contexte, ils peuvent être différents d'un cluster à un autre. Nous posons $w_j = 1$ pour tout $j \in \{1, \dots, m\}$. Pour le calcul de ζ_{jk}^h , nous posons $\tau_0 = 0.5$ et $\tau_1 = 0.75$. Les tables suivantes ont été obtenues en utilisant le logiciel R¹. Notons que pour cette illustration, nous avons opté pour un cas simple où $\mathbf{T}[\mathcal{R}]$ comme suite :

$$\mathbf{T}[\mathcal{R}] = \{j \in \mathcal{J} :: \langle 0 \leq \xi(\beta_{jk}) \leq 1 \rangle\} \quad (4.23)$$

4.4.1 Construction des clusters diagnostics \mathbf{D}_h

Le cluster diagnostic est dérivé en utilisant des trois indices (admissibilité, éligibilité et rejet) et donné par la Table 4.1 comme suit :

Remarquons que, les colonnes $\mathbf{C}_1.\mathbf{E}$ et $\mathbf{C}_3.\mathbf{E}$ n'ont aucune valeur. Cela signifie que les clusters L_1 et L_3 sont 'fortement' décrits uniquement par les modalités 'Admis'. Par contre, le cluster L_2 est 'faiblement' décrits par la modalité X_7 .Be du 7ème attribut.

4.4.2 Calcul des indices $\mathbf{A}(\mathbf{x}_i, \mathbf{D}_h)$, $\mathbf{E}(\mathbf{x}_i, \mathbf{D}_h)$ et $\mathbf{R}(\mathbf{x}_i, \mathbf{D}_h)$

Le calcul de l'indice d'admissibilité $\mathbf{A}(\mathbf{x}_i, \mathbf{D}_h)$, l'indice d'éligibilité $\mathbf{E}(\mathbf{x}_i, \mathbf{D}_h)$ et l'indice de rejet $\mathbf{R}(\mathbf{x}_i, \mathbf{D}_h)$ pour $\mathbf{x}_i \in \mathbf{U}$ et $h \in \{1, 2, 3\}$ fournit les résultats donnés par la Table (4.2). On remarque que les ménages sont affectés dans un cluster approprié pour un indice d'admissibilité plus grand que l'indice d'éligibilité correspondant. L'écart entre les deux est un indicateur de la qualité d'une affectation. Ainsi, une meilleure affectation est donnée par un indice d'admissibilité *très élevé* (proche de 1) avec un indice d'éligibilité et de rejet nul égal à zéro. Par exemple, les ménages \mathbf{x}_1 , \mathbf{x}_4 et \mathbf{x}_8 dans la Table 4.2 sont considérés comme des meilleures affectations.

1. Free downloadable on <http://www.r-project.org> [see R Development Core Team, 2011]

D ₁			D ₂			D ₃		
C ₁ .A	C ₁ .E	C ₁ .R	C ₂ .A	C ₂ .E	C ₂ .R	C ₃ .A	C ₃ .E	C ₃ .R
X ₁ .I ₁		X ₁ .I ₂	X ₁ .I ₃	X ₇ .Be	X ₁ .I ₁	X ₁ .I ₂		X ₁ .I ₁
X ₂ .N		X ₁ .I ₃	X ₂ .Y		X ₁ .I ₂	X ₂ .Y		X ₁ .I ₃
X ₃ .S		X ₂ .Y	X ₃ .V		X ₂ .N	X ₃ .A		X ₂ .N
X ₄ .N		X ₃ .A	X ₄ .Y		X ₃ .A	X ₄ .Y		X ₃ .S
X ₅ .F		X ₃ .V	X ₅ .V		X ₃ .S	X ₅ .A		X ₃ .V
X ₆ .A		X ₄ .Y	X ₆ .N		X ₄ .N	X ₆ .S		X ₄ .N
X ₇ .W		X ₅ .A			X ₅ .A			X ₅ .F
		X ₅ .V			X ₅ .F			X ₅ .V
		X ₆ .N			X ₆ .A			X ₆ .A
		X ₆ .S			X ₆ .S			X ₆ .N
		X ₇ .Be			X ₇ .U			X ₇ .Be
		X ₇ .U			X ₇ .W			X ₇ .U
								X ₇ .W

TABLE 4.1 – Clusters diagnostics

Households	Admissibility Indices			Eligibility Indices			Rejection Indices		
	C ₁	C ₂	C ₃	C ₁	C ₂	C ₃	C ₁	C ₂	C ₃
x ₁	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
x ₂	0.13	0.74	0.28	0.00	0.08	0.00	0.00	0.00	0.00
x ₃	0.27	0.25	0.59	0.00	0.00	0.00	0.02	0.03	0.01
x ₄	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
x ₅	0.84	0.00	0.16	0.00	0.00	0.00	0.01	0.00	0.01
x ₆	0.13	0.74	0.28	0.00	0.08	0.00	0.00	0.00	0.00
x ₇	0.13	0.25	0.73	0.00	0.00	0.00	0.02	0.03	0.00
x ₈	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
x ₉	0.13	0.25	0.73	0.00	0.08	0.00	0.02	0.00	0.00
x ₁₀	0.13	0.74	0.28	0.00	0.00	0.00	0.00	0.03	0.00
x ₁₁	0.84	0.00	0.16	0.00	0.00	0.00	0.01	0.00	0.01
x ₁₂	0.13	0.25	0.73	0.00	0.08	0.00	0.02	0.00	0.00

TABLE 4.2 – Les indices d’admissibilité, d’élégibilité et de rejet

4.4.3 Calcul de indices de concordance $H(x_i, D_h)$

L’indice de concordance est donné par la Table 4.3, laquelle montre que les ménages x_1 , x_4 et x_8 présentent de meilleurs indices de concordance sur \hat{C}_1 , tandis que le ménage x_3 présente un indice de concordance conflictuel sur tous les clusters.

Ménages	C_1	C_2	C_3
x_1	0.97	0.00	0.00
x_2	0.13	0.82	0.28
x_3	0.27	0.25	0.59
x_4	0.97	0.00	0.00
x_5	0.84	0.00	0.16
x_6	0.13	0.82	0.28
x_7	0.13	0.25	0.73
x_8	0.97	0.00	0.00
x_9	0.13	0.33	0.73
x_{10}	0.13	0.74	0.28
x_{11}	0.84	0.00	0.16
x_{12}	0.13	0.33	0.73

TABLE 4.3 – Les indices de concordance

4.4.4 Calcul des indices de discordance $D(x_i, D_h)$

L'indice de discordance est donné par la Table 4.4 laquelle montre que les ménages x_1 , x_4 et x_8 présentent des indices de discordance nuls sur C_1 , tandis que le ménage x_3 présente un indice de discordance conflictuel sur tous les clusters.

Households	C_1	C_2	C_3
x_1	0.00	1.00	1.00
x_2	0.96	0.14	0.80
x_3	0.84	0.87	0.38
x_4	0.00	1.00	1.00
x_5	0.12	1.00	0.94
x_6	0.96	0.14	0.80
x_7	0.97	0.87	0.24
x_8	0.00	1.00	1.00
x_9	0.97	0.86	0.24
x_{10}	0.96	0.25	0.80
x_{11}	0.12	1.00	0.94
x_{12}	0.97	0.86	0.24

TABLE 4.4 – Les indices de discordance

4.4.5 Calcul des indices d'appartenance $B(x_i, D_h)$

Une affectation crispée donnée par la Table 4.5 nous conduit à la conclusion donnée par la Table 4.6. Remarquons que x_3 est plus faiblement affecté à la capa-

Households	C_1	C_2	C_3
x_1	0.97	0.00	0.00
x_2	0.04	0.82	0.20
x_3	0.16	0.13	0.59
x_4	0.97	0.00	0.00
x_5	0.84	0.00	0.06
x_6	0.04	0.82	0.20
x_7	0.03	0.13	0.73
x_8	0.97	0.00	0.00
x_9	0.03	0.14	0.73
x_{10}	0.04	0.74	0.20
x_{11}	0.84	0.00	0.06
x_{12}	0.03	0.14	0.73

TABLE 4.5 – Affectations ‘crispées’

Capabilités Étendues	Degré d’appartenance des ménages
C_1	$(x_1 :0.969)$; $(x_4 :0.969)$; $(x_5 :0.842)$; $(x_8 :0.969)$; $(x_{11} :0.842)$.
C_2	$(x_2 :0.82)$; $(x_6 :0.82)$; $(x_{10} :0.737)$.
C_3	$(x_3 :0.592)$; $(x_7 :0.726)$; $(x_9 :0.726)$; $(x_{12} :0.726)$.

TABLE 4.6 – Affectations floues

bilité étendue C_3 que dans les autres affectations. Cela est dû au fait que les modalités ‘F’ et ‘U’ décrivent *faiblement* x_3 dans le cluster L_3 . Toutefois, La Table (4.6) conclut à 100% les affectations correctes.

Ce chapitre 4 a consisté à la présentation de la deuxième phase de notre méthodologie. En résumé, cette phase vise à la formulation d’une procédure d’affectation de ‘nouveaux’ ménages de l’univers U (la population totale) à une ou plusieurs classes en examinant les vecteurs de commodités des ménages et en se référant aux clusters diagnostics, aux indices d’admissibilité, aux indices d’éligibilité et aux indices de rejet que nous avons largement développé dans ce chapitre. Nous supposons que les classes ne sont pas ordonnées. Ils sont plutôt décrits/caractérisés par une ou plusieurs ‘modalités de référence’ regroupées en sous-ensembles tel que chaque sous-ensemble, appelé ‘*cluster diagnostic*’, soit associé à sa classe correspondante. Delà nous affectons les nouveaux ménages aux classes spécifiques sur la base des clusters diagnostics. Pour atteindre cet objectif, nous avons proposé une nouvelle méthode de filtrage multi-attributs basée sur les principes de concordance et de discordance dérivés des concepts bien–connus de majorité et de veto.

Conclusion générale

Nous avons introduit dans cette thèse une nouvelle méthodologie pour l'analyse et le mesurage multidimensionnelle de la pauvreté en se basant sur l'approche par les capacités de Sen sous une perspective d'aide à la décision. Initialement, notre méthodologie commence par une sélection des individus et une sélection des variables sur lesquelles la segmentation doit être effectuée de manière à discriminer autant que possible les clusters et mettre en lumière les informations pertinentes pour notre situation problématique. La table de données considérée en entrée, représentant un ensemble de n vecteurs de commodités, est l'évaluation d'un échantillon de n ménages sur m variables mixtes. Cette table de données est organisée en clusters de telle sorte que chaque cluster soit formé des ménages similaires entre eux en termes de distribution des commodités et différents de ménages des autres clusters.

Nous établissons les dimensions de bien-être et des niveaux des dimensions de bien-être à partir desquels nous construisons les fonctionnements, puis les capacités réalisables correspondant à chaque cluster. Chaque capacité réalisable, associée à un cluster donné, reflète l'aptitude ou l'habileté des ménages de ce cluster à accomplir ou à réaliser un sous-ensemble de fonctionnements parmi les différents faisceaux de fonctionnements alternatifs. Selon les commodités génériques, la capacité réalisable d'un cluster peut être améliorée de manière à obtenir sa capacité étendue. Cela permet de prendre en compte, lors de la construction des mesures significatives de la pauvreté (MDPM), certaines commodités telles que le revenu, le statut professionnel, le statut matrimonial, l'âge du chef de ménage, la taille du ménage, etc. Plus précisément, une MDPM est un instrument qui aide le décideur à prendre des décisions bien-informées au sujet des politiques, programmes et projets en mettant les meilleures évidences de la recherche au cœur de l'élaboration et de l'implémentation des politiques. Ainsi, elle permet aux décideurs de cibler les ressources et de concevoir des politiques plus efficaces pour l'allocation ou la redistribution de ces ressources. Cette nouvelle façon d'aborder la problématique du mesurage de la pauvreté permet d'évaluer la pauvreté à partir de ce que les ménages (ou les individus) sont *capables de faire* et d'*être* au vue de leurs vecteurs de commodités. Ce qui fait de cet instrument un outil utile pour le suivi régulier et le diagnostic des problèmes sociaux, lequel est également utilisable pour le contrôle et l'évaluation des politiques d'inclusion sociale.

D'autre part, nous avons introduit une méthode de filtrage multi-attributs permettant d'affecter des nouveaux ménages au cluster adéquat. Dans cette phase, la table de données "complète" considérée en entrée, représentant un ensemble de n vecteurs de commodités, est l'évaluation de n ménages de l'univers sur m variables mixtes. Notre méthode d'affectation se base sur le concept de cluster diagnostic que

nous avons développé dans cette thèse. Les clusters diagnostics sont le résultat de la transformation de la table des données de grande taille en un tableau récapitulatif dans le but d'acquérir des connaissances cruciales sur chaque cluster. C'est un outil visant à identifier et à décrire la nature et les raisons des différents types de pauvreté observés au sein de la population échantillon de manière à produire un diagnostic approprié de la situation. Ainsi, la méthode d'affectation proposée dans cette thèse fournit au décideur une information nuancée et fiable sur le degré d'appartenance de chaque ménage de l'univers U (la population totale) aux différents clusters. Les indices d'admissibilité, d'éligibilité et de rejet introduits dans cette thèse sont cruciaux pour analyser et trancher par rapport à la solidité ou la faiblesse de l'affectation d'un ménage dans un cluster donné. En outre, ces indices peuvent être directement utilisés comme variables d'entrée par un module d'exécution à des fins d'automatisation.

La méthodologie proposée dans cette thèse peut donc être utilisée pour établir des moyens rigoureux de définition et d'évaluation des coûts et résultats d'une politique donnée. C'est également un moyen de promouvoir la transparence dans la gestion des ressources par le biais de l'étude coût-efficacité, coût-bénéfice, coût-utilité des politiques, programmes ou projets. Notre méthodologie propose ainsi des mesures significatives satisfaisant trois compétences fondamentales : *compétence cognitive, compétence analytique, compétence prédiction des conséquences et évaluation des résultats*.

En guise de perspectives, notre méthodologie peut être étendue à des situations plus complexes pour mettre en évidence la dynamique de la pauvreté au sein des sociétés des pays en développement, ainsi qu'une image vivante de différents types de pauvreté multidimensionnelle, à la fois à travers les pays, les régions et dans le monde ; et dans les pays à travers les groupes ethniques, les zones urbaine/rurale, ou d'autres caractéristiques clés. Nous comptons également étudier l'aspect axiomatique des indices développés dans cette thèse.

Application sur la base de données ASSL 2007 du Burkina Faso

A.1 Présentation de la base des données

Le processus de validation de notre approche opérationnelle exige que nous ayons en entrée une base de données d'enquête sur les conditions de vie. L'Institut National de la Statistique et de la Démographie (INSD) du Burkina Faso conduit depuis 1994 pour le compte et à la demande du gouvernement burkinabè des enquêtes d'envergure nationale sur les conditions de vie des ménages. La réalisation de ces enquêtes s'inscrivent dans la perspective d'une meilleure compréhension de la pauvreté au Burkina Faso, pour un meilleur suivi de celle-ci à travers ses diverses manifestations. Elle vise à fournir des données utiles à l'affinement de l'analyse au sein des différents groupes sectoriels et thématiques du dispositif institutionnel de suivi de la mise œuvre du cadre stratégique de lutte contre la pauvreté (CSLP).

La base de données choisit dans le cadre de cette thèse pour des besoins d'illustration porte sur l'enquête annuelle sur les conditions de vie des ménages (EA/QUIBB) réalisée en 2007 par l'INSD. Elle nous permettra *in fine* de valider notre modèle et d'obtenir des éléments pour appréhender les différentes pauvretés au Burkina Faso (particulièrement dans la ville de Ouagadougou) afin de proposer des politiques de lutte adéquates. Cette base de données permettra également d'apprécier les tendances d'évolution de la pauvreté au Burkina Faso par la mise à jour des indicateurs. Un indicateur est une mesure qui nous aide à comprendre où nous en sommes, où nous allons et combien nous sommes loin de l'objectif. Il permet de résumer les caractéristiques des systèmes ou de mettre en évidence ce qui se passe dans un système. Nous faisons l'hypothèse dans cette thèse que les variables décrivent les commodités des ménages.

Le Questionnaire Unifié des Indicateurs de Base du Bien-être (QUIBB) est une enquête auprès des ménages du Burkina Faso dont l'un des objectifs est de mieux connaître leurs conditions de vie et de constater les changements qui s'opèrent dans l'espace et le temps en ce qui concerne leur bien-être. Tenant compte de l'importance de cette problématique, les enquêtes annuelles sur les conditions de vie des ménages (EA/QUIBB) accordent une attention particulière à l'appréhension des caractéristiques des ménages en général et à celles des chefs de ménage en particu-

12 Annexe A. Application sur la base de données ASSL 2007 du Burkina Faso

lier. Les conditions de vie des ménages concernent la satisfaction ou non de leurs besoins essentiels : nourriture, logement, santé, éducation, loisirs, etc. Les déterminants des conditions dans lesquelles vivent les membres du ménage dépendent pour une large part des caractéristiques démographiques et socio-économiques mêmes du chef de ménage dont le sexe, l'âge, l'état matrimonial, le niveau d'instruction, le handicap et le groupe socio-économique. Pour ce faire, le chef de ménage joue un rôle de premier plan au sein du ménage. D'autres paramètres comme la taille du ménage, le milieu de résidence (rural, urbain) et les activités économiques exercées par les différents membres, déterminent aussi, à des degrés divers, les conditions de vie du ménage.

L'ensemble des données d'enquête (1255 ménages sur 48 variables) utilisé afin d'illustrer notre méthodologie est un échantillon national représentatif de ménages vivant à Ouagadougou (la capitale du Burkina Faso). Le QUIBB fournit des informations sur les dimensions de la pauvreté que nous avons regroupés en six dimensions de bien-être caractérisant un cluster L_h de la population Ω : la nutrition, l'éducation, l'eau et l'assainissement, le logement, la santé et le transport. Notons que, la décision concernant le nombre de dimension de bien-être potentiel et le choix des variables inhérentes dépendent du décideur et du contexte dans lequel l'étude est menée. Les choix que nous avons fait dans le présent document concernant les six dimensions de bien-être sont uniquement pour des raisons d'illustration.

Remarque A.1.1 *Cette enquête a été réalisée dans le cadre des Objectifs du Millénaire pour le Développement (OMD). Les OMD et les cibles proviennent de la Déclaration du Millénaire, signée par 189 pays, dont 147 chefs d'Etat et de gouvernement, en Septembre 2000 [voir [UNDP, 2003](#)].*

- **Nutrition** : Le QUIBB fournit l'information nutritionnelle pour chaque membre du ménage. Nous utilisons trois dimensions de la pauvreté afin de déterminer si un ménage est privé ou vulnérables en terme de nutrition (mal-nourris) en mesurant directement si un ménage a eu des difficultés à satisfaire ses besoins fondamentaux et sa situation économique.
- **Education** : Le QUIBB fournit des informations sur le niveau scolaire de chaque membre du ménage. Le nombre de scolarité capture le niveau de connaissance et de compréhension de chaque ménage. Tout comme le niveau scolaire, il ne tient pas compte de la qualité de l'éducation, ni du niveau de compétences. C'est pourquoi dans cette thèse, nous avons considéré comme indicateur relatif le niveau du fonctionnement.
- **Eau et Assainissement (WS)** : Le facteur *Eau et Assainissement* utilise six dimensions de pauvreté, lesquelles représentent le niveau de déprivation sur chaque cluster de ménages en termes d'accès à l'eau potable et à un assainissement adéquat. Il comprend deux indicateurs des OMD standard (eau

potable et assainissement amélioré) qui fournissent quelques indications rudimentaires de la qualité des services d'eau et d'assainissement pour les ménages.

- **Logement** : Le facteur *Logement* utilise 21 dimensions de pauvreté, lesquelles représentent le niveau de déprivation en terme de logement sur chaque cluster de ménages. Il comprend un indicateur des OMD standard (l'utilisation de combustible propre pour la cuisine) et 20 indicateurs non-OMD (tels que l'électricité, le revêtement du sol). Deux d'entre eux fournissent quelques indications rudimentaires de la qualité des logements pour chaque cluster.
- **Santé** : Le facteur *Santé* utilise une dimension de la pauvreté qui évalue l'accès aux services de santé indépendamment du mode de transport nécessaire pour y accéder.
- **Mobilité** : Ce dernier facteur couvre la possession de certains biens de consommation pour le transport telles que vélo, en moto, voiture et l'accès à un service de transport public.

Remarquons que, chaque $F_{\mathcal{J}_u}$, $u \in \{1, 2, 3, 4, 5, 6\}$ est perçue comme l'ensemble de toutes les commodités associées à l'ensemble $\prod_{j \in \mathcal{J}_u} X_j$ dans population totale Ω et $F_{\mathcal{J}_u}(L_h)$ est perçue comme l'ensemble de toutes les commodités associées à l'ensemble $\prod_{j \in \mathcal{J}_u} X_j$ dans le cluster L_h .

Le problème commun de données manquantes peut survenir durant la classification. Pour résoudre ce problème, nous avons uniquement traité les dimensions contenant des valeurs valides. Ainsi, cette méthodologie peut également être appliquée à d'autres bases de données telles que les Enquêtes démographiques et de santé (EDS) et Enquêtes de l'organisation mondiale de la santé (OMS). Notre choix pour EA/QUIBB 2007 a été motivé uniquement par des besoins d'illustration.

A.2 Application

Soit l'EA/QUIBB de 1255 ménages évalués sur 48 variables et six dimensions potentielles de bien-être. Nous avons effectué une classification non supervisée via bootstrap multi-échelle (nombre de bootstrap 1000 ; [voir Suzuki et Shimodaira, 2006]) en utilisant la méthode de Ward [Ward, 1963] et une matrice de corrélation basée sur la dissimilarité. Le meilleur nombre de clusters obtenu est de huit clusters. Ce résultat peut également être confirmé par l'indice de Calinski et Harabasz [1974], lequel permet de comparer l'homogénéité des partitions. En utilisant le logiciel R avec $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8) = (0.063, 0.272, 0.224, 0.164, 0.09, 0.118, 0.028, 0.042)$ et $\mathbf{L} = \{L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8\}$, nous obtenons les tables suivantes.

À partir de l'équation (3.29), nous obtenons la représentation ordonnée (A.1)

12 Annexe A. Application sur la base de données ASSL 2007 du Burkina Faso

Clusters	Logement		Mobilité		Nutrition		WS		Santé		Education	
	$S_{\mathcal{J}_1}$	$U_{\mathcal{J}_1}$	$S_{\mathcal{J}_2}$	$U_{\mathcal{J}_2}$	$S_{\mathcal{J}_3}$	$U_{\mathcal{J}_3}$	$S_{\mathcal{J}_4}$	$U_{\mathcal{J}_4}$	$S_{\mathcal{J}_5}$	$U_{\mathcal{J}_5}$	$S_{\mathcal{J}_6}$	$U_{\mathcal{J}_6}$
C_1	0.87	0.04	0.82	0.05	0.92	0.07	0.87	0.07	0.72	0.08	0.86	0.11
C_2	0.16	0.95	0.24	0.67	0.53	0.59	0.18	1.00	0.26	0.63	0.37	0.78
C_3	0.37	0.72	0.70	0.41	0.66	0.36	0.44	0.78	0.45	0.35	0.38	0.51
C_4	0.11	1.00	0.08	0.81	0.42	0.54	0.08	0.80	0.28	0.35	0.21	0.65
C_5	0.74	0.10	0.74	0.06	0.84	0.16	0.82	0.09	0.86	0.03	0.77	0.16
C_6	0.60	0.22	0.62	0.17	0.84	0.14	0.72	0.24	0.89	0.05	0.63	0.22
C_7	0.99	0.05	1.00	0.03	1.00	0.15	0.88	0.04	1.00	0.00	1.00	0.01
C_8	0.94	0.03	0.91	0.02	0.69	0.15	0.98	0.02	0.87	0.01	0.93	0.02

TABLE A.1 – Evaluation de $C_h(S_{\mathcal{J}_u})$ et $C_h(U_{\mathcal{J}_u})$

qui montre que la capacité réalisable C_4 est la plus pauvre, donc la plus prioritaire en terme de politiques d'intervention ou d'implémentation de politiques.

$$C_7 \succ C_8 \succ C_1 \succ C_5 \succ C_6 \succ C_3 \succ C_2 \succ C_4 \quad (\text{A.1})$$

À partir de la Table A.2, nous pouvons dériver les *seuils de fonctionnements* $F_{\text{sup}}(L_h)$ lesquels représentent les premiers vecteurs de fonctionnements dans l'ensemble des capacités de la Table A.3. Nous avons par exemple, $F_{\text{sup}}(L_1) = \langle G, G, G, G, A, G \rangle$ and $F_{\text{sup}}(L_2) = F_{\text{sup}}(L_4) = \langle B, B, B, B, B, B \rangle$, $F_{\text{sup}}(L_3) = \langle B, A, A, B, B, B \rangle$, $F_{\text{sup}}(L_5) = F_{\text{sup}}(L_6) = \langle A, A, A, A, G, A \rangle$, $F_{\text{sup}}(L_7) = \langle G, G, G, G, G, G \rangle$, $F_{\text{sup}}(L_8) = \langle G, G, A, G, G, G \rangle$.

	\mathcal{J}_1	ε_1	\mathcal{J}_2	ε_2	\mathcal{J}_3	ε_3	\mathcal{J}_4	ε_4	\mathcal{J}_5	ε_5	\mathcal{J}_6	ε_6
M_1	0.84	0.96	0.78	0.95	0.85	0.93	0.80	0.93	0.65	0.90	0.75	0.89
M_2	-0.78	0.85	-0.43	0.74	-0.06	0.53	-0.82	0.85	-0.36	0.70	-0.41	0.68
M_3	-0.35	0.66	0.29	0.63	0.29	0.64	-0.34	0.64	0.10	0.56	-0.13	0.57
M_4	-0.89	0.90	-0.73	0.91	-0.11	0.56	-0.72	0.91	-0.06	0.55	-0.44	0.76
M_5	0.64	0.88	0.68	0.93	0.68	0.84	0.73	0.90	0.84	0.97	0.61	0.83
M_6	0.38	0.73	0.45	0.79	0.70	0.85	0.48	0.75	0.84	0.94	0.42	0.74
M_7	0.94	0.95	0.97	0.97	0.85	0.87	0.84	0.95	1.00	1.00	0.99	0.99
M_8	0.92	0.97	0.89	0.98	0.55	0.83	0.96	0.98	0.86	0.99	0.90	0.98

TABLE A.2 – Les mesures significantes

Toutefois, une analyse rapide des clusters montre que les 'pires' capacités réalisables (exemple $\{C_3; C_2; C_4\}$) sont caractérisés par le fait que le chef de ménage est *Agriculteur* et le 'meilleures' (exemple $\{C_7; C_8; C_1\}$) sont caractérisées par le fait que le chef de ménage est *Employé de Service Public*. Il est important de noter que, les mots 'pire' et 'meilleure' ne contiennent rien de numérique, mais ils reflètent uniquement les opportunités et le bien-être des ménage d'un cluster donné.

$$\begin{aligned}
C_1 &= \{ \langle G, G, G, G, A, G \rangle; \langle G, G, G, G, A, A \rangle; \langle G, G, G, A, A, A \rangle; \dots \langle B, B, B, B, B, B \rangle \} \\
C_2 &= \{ \langle B, B, B, B, B, B \rangle \} \\
C_3 &= \{ \langle B, A, A, B, B, B \rangle; \langle B, B, A, B, B, B \rangle; \langle B, A, B, B, B, B \rangle; \langle B, B, B, B, B, B \rangle \} \\
C_4 &= \{ \langle B, B, B, B, B, B \rangle \} \\
C_5 &= \{ \langle A, A, A, A, G, A \rangle; \langle A, A, A, A, A, A \rangle; \langle A, A, A, A, A, B \rangle; \dots \langle B, B, B, B, B, B \rangle \} \\
C_6 &= \{ \langle A, A, A, A, G, A \rangle; \langle A, A, A, A, A, A \rangle; \langle A, A, A, A, A, B \rangle; \dots \langle B, B, B, B, B, B \rangle \} \\
C_7 &= \{ \langle G, G, G, G, G, G \rangle; \langle G, G, G, G, A, G \rangle; \langle G, G, G, G, G, A \rangle; \dots \langle B, B, B, B, B, B \rangle \} \\
C_8 &= \{ \langle G, G, A, G, G, G \rangle; \langle G, G, A, G, G, A \rangle; \langle G, G, A, A, G, G \rangle; \dots \langle B, B, B, B, B, B \rangle \}
\end{aligned}$$

TABLE A.3 – Ensembles des capacités réalisables

Considérons maintenant la commodité générique ‘revenu’¹ et les capacités réalisables C_5 et C_6 (voir Table A.3) dont nous aimerions augmenter certains des niveaux réalisables du vecteur fonctionnement. Une observation rapide montre que les seuils de fonctionnement $F_{\text{sup}}(L_5) = F_{\text{sup}}(L_6)$. On peut être tenté de conclure que “Le niveau de vie (bien-être) offerts par C_5 est considéré être équivalent au niveau de vie (ou bien-être) offert par C_6 ”. Toutefois, la Table A.2 montre que seules $M_5(\mathcal{J}_5)$ et $M_6(\mathcal{J}_5)$ sont égales en terme de mesures significatives. De plus, la Table A.5 montre que $X_2(L_5) = [90000, 110000]$, tandis que $X_2(L_6) = [63000, 90000]$. Ainsi, nous pouvons intuitivement augmenter les niveaux de la dimension de bien-être \mathcal{J}_3 de A à G, c’est-à-dire où $M_5(\mathcal{J}_3) < M_6(\mathcal{J}_3)$ in Table A.2. Les capacités étendues des capacité réalisable C_5 and C_6 sont alors données par les ensembles décrits dans (A.2) et (A.3).

$$\tilde{C}_5 = \{ \langle A, A, G, A, G, A \rangle; \langle A, A, G, A, A, A \rangle; \dots \} \quad (\text{A.2})$$

$$\tilde{C}_6 = \{ \langle A, A, A, A, G, A \rangle; \langle A, A, A, A, A, A \rangle; \dots \} \quad (\text{A.3})$$

On suppose que le décideur a imposé des niveaux de contrôle de telle sorte que les capacités étendues des capacités réalisables C_7 et C_8 soient données par

$$\tilde{C}_7 = \{ \langle G, G, G, G, G, G \rangle; \langle G, G, G, A, G, G \rangle; \langle G, G, A, G, G, G \rangle \}$$

et

$$\tilde{C}_8 = \{ \langle G, G, A, G, G, G \rangle; \langle G, G, A, G, G, A \rangle; \langle G, G, A, A, G, G \rangle \}.$$

Ainsi, les capacités étendues sont données comme suit :

1. Notons que le ‘revenu’ dans l’EA/QUIBB 2007 est donné en CFA franc. Le franc CFA est la monnaie de plusieurs pays d’Afrique constituant en partie la Zone franc : 1 euro = 655.957 CFA.

$$\begin{aligned} \tilde{C}_1 &= \{ \langle G, G, G, G, A, G \rangle; \langle G, G, G, G, A, A \rangle; \dots \langle B, B, B, B, B, B \rangle \} \\ \tilde{C}_2 &= \{ \langle B, B, B, B, B, B \rangle \} \\ \tilde{C}_3 &= \{ \langle B, A, A, B, B, B \rangle; \langle B, B, A, B, B, B \rangle; \langle B, A, B, B, B, B \rangle; \langle B, B, B, B, B, B \rangle \} \\ \tilde{C}_4 &= \{ \langle B, B, B, B, B, B \rangle \} \\ \tilde{C}_5 &= \{ \langle A, A, G, A, G, A \rangle; \langle A, A, G, A, A, A \rangle; \langle A, A, A, A, G, A \rangle; \dots \} \\ \tilde{C}_6 &= \{ \langle A, A, A, A, G, A \rangle; \langle A, A, A, A, A, A \rangle; \langle A, A, A, A, A, B \rangle; \dots \} \\ \tilde{C}_7 &= \{ \langle G, G, G, G, G, G \rangle; \langle G, G, G, A, G, G \rangle; \langle G, G, A, G, G, G \rangle \} \\ \tilde{C}_8 &= \{ \langle G, G, A, G, G, G \rangle; \langle G, G, A, G, G, A \rangle; \langle G, G, A, A, G, G \rangle \} \end{aligned}$$

TABLE A.4 – Ensembles des capacités étendues

	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈
	\mathcal{X}_2 : Revenu en CFA (Échelle 1/10000)							
	[12, 15]	[1.75, 5]	[3.5, 6.5]	[0.1, 2.5]	[9, 11]	[6.3, 9]	[25, 60]	[17, 22.5]
	\mathcal{X}_1 : Taille des ménages							
[1, 3]	0.266	0.534	0.406	0.539	0.319	0.318	0.343	0.288
[4, 5]	0.316	0.252	0.31	0.262	0.336	0.345	0.343	0.308
[6, 30]	0.418	0.214	0.285	0.199	0.345	0.338	0.314	0.404
	\mathcal{X}_3 : Âge (en années)							
[16, 34]	0.241	0.443	0.349	0.374	0.23	0.324	0.114	0.173
[35, 47]	0.354	0.323	0.324	0.301	0.372	0.338	0.314	0.365
[48, 99]	0.405	0.235	0.327	0.325	0.398	0.338	0.571	0.462

TABLE A.5 – Description de quelques commodités génériques

A.2.1 Construction des clusters diagnostics D_h

Le cluster diagnostic d'une base de données EA/QUIBB est donné selon trois niveaux de "force" par la Table A.6. Nous pouvons dériver les conclusions suivantes après une observation rapide :

- \tilde{C}_2 , \tilde{C}_3 et \tilde{C}_4 représentent l'ensemble des ménages les plus "pauvres" entre eux (en termes de déprivation sur certaines variables) par rapport aux ménages d'autres capacités étendues. Mais, \tilde{C}_2 est plus proche (en termes de déprivations) de \tilde{C}_4 que \tilde{C}_3 . Elles représentent les clusters de ménages qui sont déprivés sur la plupart des variables. Toutefois, il existe une 'légère' différence sur la façon dont ils sont déprivés. Par exemple, \tilde{C}_4 est déprivé sur plus de variables que \tilde{C}_2 tel que nous pouvons le voir à partir de $X_{10}.N$, $X_{15}.N$, $X_{19}.N$ dans ces deux capacités étendues. Tandis que les ménages de \tilde{C}_3 sont en majorité non-déprivés sur les variables $X_{16}.Y$, $X_{22}.Y$, $X_{30}.Y$, $X_{31}.Y$; sur la variable $X_{15}.Y$, ils sont déprivés (respectivement, partiellement déprivés) dans \tilde{C}_4 (respectivement dans \tilde{C}_2). Nous pouvons trivialement noter que après l'affectation floue, les ménages dans \tilde{C}_4 sont aussi inclus dans

- \tilde{C}_2 mais l'inverse n'est pas vrai.
- Les ménages de \tilde{C}_7 sont non-déprivés sur la majorité des variables que ceux de toutes les autres capacités étendues existantes. Toutefois, nous pouvons remarquer que les ménages de \tilde{C}_8 et \tilde{C}_1 sont moins déprivés et plus proches de C_7 que tous les ménages ne le sont. Nous pouvons également noter quelques modalités (tels que $X_{12}.Y$, $X_{24}.V$ et $X_{33}.V$) qui sont présents dans \tilde{C}_7 comme 'Admis' et dans \tilde{C}_8 et C_1 comme 'Eligible'.
 - Les ménages de \tilde{C}_5 et \tilde{C}_6 sont non-déprivés dans certaines variables et déprivés dans d'autres. Néanmoins, les ménages de \tilde{C}_5 présentent une variable majoritairement déprivée ($X_{34}.N$) comme 'Admis' laquelle apparaît aussi dans \tilde{C}_6 comme 'Admitted', et une variable déprivée ($X_{10}.N$) comme 'Eligible' laquelle n'apparaît pas dans \tilde{C}_6 tandis que les ménages de \tilde{C}_6 présentent les variables déprivées ($X_{34}.N$ and $X_{18}.N$) comme 'Admis'.

Une telle analyse peut être étendue à d'autres clusters ou d'autres variables de manière à appréhender les différents types de pauvreté ou de bien-être présents dans la société. Ceci est très utile pour l'élaboration des politiques et pour l'implémentation des politiques. Les clusters diagnostics sont donc un outil puissant pour l'analyse des politiques et peuvent être utiles pour la conception et l'identification d'alternatives préférables en ce qui concerne les questions de politiques complexes.

D ₁		D ₂		D ₃		D ₄		D ₅		D ₆		D ₇		D ₈	
C ₁ .A	C ₁ .E	C ₂ .A	C ₂ .E	C ₃ .A	C ₃ .E	C ₄ .A	C ₄ .E	C ₅ .A	C ₅ .E	C ₆ .A	C ₆ .E	C ₇ .A	C ₇ .E	C ₈ .A	C ₈ .E
X ₆ .M	X ₁₂ .Y	X ₁₁ .N	X ₁₀ .N	X ₁₃ .N	X ₁₂ .N	X ₁₀ .N	X ₃₉ .P	X ₇ .M	X ₈ .H	X ₆ .M	X ₁₃ .Y	X ₇ .M	X ₆ .M	X ₇ .M	X ₆ .M
X ₇ .M	X ₁₃ .Y	X ₁₂ .N	X ₁₅ .N	X ₁₄ .N	X ₁₅ .Y	X ₁₁ .N	X ₄₀ .O	X ₁₅ .Y	X ₁₀ .N	X ₇ .M	X ₃₃ .V	X ₈ .H	X ₁₈ .Y	X ₈ .H	X ₁₁ .Y
X ₈ .H	X ₂₄ .V	X ₁₃ .N	X ₁₉ .N	X ₁₆ .Y	X ₃₉ .P	X ₁₂ .N	X ₄₈ .N	X ₁₆ .Y	X ₁₄ .Y	X ₁₅ .Y	X ₄₅ .V	X ₁₂ .Y	X ₃₀ .Y	X ₁₃ .Y	X ₁₂ .Y
X ₁₄ .Y	X ₃₃ .V	X ₁₄ .N	X ₂₆ .E	X ₁₈ .N		X ₁₃ .N		X ₁₇ .Y	X ₂₀ .Y	X ₁₆ .Y	X ₄₇ .V	X ₁₃ .Y	X ₄₄ .V	X ₁₄ .Y	X ₂₄ .V
X ₁₅ .Y	X ₃₈ .V	X ₁₇ .N	X ₂₉ .K	X ₂₀ .N		X ₁₄ .N		X ₁₉ .Y	X ₃₃ .V	X ₁₇ .Y		X ₁₄ .Y		X ₁₅ .Y	X ₂₅ .T
X ₁₆ .Y	X ₄₅ .V	X ₁₈ .N		X ₂₁ .N		X ₁₅ .N		X ₂₂ .Y	X ₃₈ .V	X ₁₈ .N		X ₁₅ .Y		X ₁₆ .Y	X ₃₃ .V
X ₁₇ .Y		X ₂₀ .N		X ₂₂ .Y		X ₁₇ .N		X ₂₃ .D	X ₃₉ .I	X ₁₉ .Y		X ₁₆ .Y		X ₁₇ .Y	X ₄₂ .P
X ₁₉ .Y		X ₂₁ .N		X ₂₄ .S		X ₁₈ .N		X ₂₆ .C	X ₄₆ .V	X ₂₂ .Y		X ₁₇ .Y		X ₁₉ .Y	X ₄₅ .V
X ₂₀ .Y		X ₂₃ .D		X ₂₅ .C		X ₁₉ .N		X ₂₉ .E	X ₄₇ .V	X ₂₆ .C		X ₁₉ .Y		X ₂₀ .Y	X ₄₈ .H
X ₂₂ .Y		X ₂₄ .S		X ₂₆ .C		X ₂₀ .N		X ₃₁ .Y		X ₂₉ .E		X ₂₀ .Y		X ₂₁ .Y	
X ₂₃ .D		X ₂₅ .C		X ₂₇ .M		X ₂₁ .N		X ₃₄ .N		X ₃₁ .Y		X ₂₁ .Y		X ₂₂ .Y	
X ₂₆ .C		X ₂₇ .M		X ₃₀ .Y		X ₂₂ .N		X ₄₁ .D		X ₃₄ .N		X ₂₂ .Y		X ₂₃ .D	
X ₂₈ .G		X ₃₂ .N		X ₃₁ .Y		X ₂₃ .D		X ₄₃ .V		X ₃₇ .I		X ₂₃ .D		X ₂₆ .C	
X ₂₉ .E		X ₃₇ .I		X ₃₂ .N		X ₂₄ .S		X ₄₄ .V		X ₄₁ .D		X ₂₄ .V		X ₂₈ .G	
X ₃₁ .Y		X ₃₉ .P		X ₄₀ .OL		X ₂₅ .C				X ₄₃ .V		X ₂₅ .T		X ₂₉ .E	
X ₃₄ .N		X ₄₀ .OL		X ₄₂ .S		X ₂₆ .E				X ₄₄ .V		X ₂₆ .C		X ₃₀ .Y	
X ₃₇ .I		X ₄₂ .S		X ₄₃ .V		X ₂₈ .FS						X ₂₈ .G		X ₃₁ .Y	
X ₃₉ .I				X ₄₄ .V		X ₂₉ .K						X ₂₉ .E		X ₃₂ .Y	
X ₄₁ .D						X ₃₁ .N						X ₃₁ .Y		X ₃₄ .N	
X ₄₃ .V						X ₃₂ .N						X ₃₂ .Y		X ₃₇ .I	
X ₄₄ .V												X ₃₃ .V		X ₃₉ .I	
X ₄₆ .V												X ₃₄ .N		X ₄₁ .D	
X ₄₇ .V												X ₃₈ .V		X ₄₃ .V	
												X ₃₉ .I		X ₄₄ .V	
												X ₄₀ .V		X ₄₆ .V	
												X ₄₁ .D		X ₄₇ .V	
												X ₄₂ .P		X ₄₈ .H	
												X ₄₅ .V			
												X ₄₆ .V			
												X ₄₇ .V			
												X ₄₈ .H			

TABLE A.6 – Ensemble de clusters diagnostics de l'EA/QUIBB

A.2.2 Résultats et discussion

Nous avons testé notre nouvelle méthode de filtrage multi-attribut dans cette thèse sur un exemple illustratif de 12 ménages donnés par la Table (3.1) et sur la base de données EA-QUIBB 2007 (1255 ménages et 48 variables). Le résultat obtenu par la nouvelle méthode de filtrage multi-attribut est de 100% d'affectation floue pour tous les cluster L_1 , L_2 and L_3 . Tandis que les pourcentages d'affectation correcte et d'affectation incorrecte pour la base de données EA-QUIBB 2007 sont résumés dans la Table (A.7) avec 76.6% de cas d'affectation correcte. Une analyse de la sensibilité montre que les cas d'affectation incorrecte sont due au fait que nous avons utilisé des clusters obtenus d'une classification non supervisées avec une probabilité 'élevée' de contenir des ménages mal-affectés dans les différents clusters. Nous appelons cela : *erreurs de propagation* d'affectations incorrectes dues à la classification non supervisée. D'autre part, on note la présence des individus atypiques durant la détermination du meilleur nombre de clusters [see [Kana et al., 2012](#)]. Ces individus atypiques sont des ménages qui sont numériquement distant du reste des données et ils peuvent se présenter dans toute la distribution avec une probabilité faible. Toutefois, ils sont souvent des indicateurs d'erreur de mesurage (les méthodes de classification ou de calcul de la matrice de dissimilarité) ou des indicateurs que la population présente une *queue de distribution* [voir [Jain et Dubes, 1988](#), pour plus de détails].

Clusters	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈
Cas affectation correcte (%)	92.41	54.55	43.06	78.16	83.19	79.73	88.57	92.31
Cas d'affectation incorrecte (%)	7.59	45.45	56.94	21.84	16.81	20.27	11.43	7.69

TABLE A.7 – Pourcentage des cas d'affectation correcte et incorrecte

X_i	Description : range
X_1	Size of Household : 0, 1, 2, . . .
X_2	Monthly Income for Basic Needs of Household : > 0
X_3	Age (in years) : ≥ 0
X_4	HhWEIGHT : Household Weight (> 0)
X_5	Occupational Status of Household Head : Public Sector Employee(Pu), Private Sector Employee(Pr), Employee of the Informal Private(I), Farmer(F), Cotton Farmer(C), Other Type of Agriculture (O), Familial help-Volunteer-Apprentice(FVA), Inactive(N), Jobless(J)
X_6	Gender : Male(M), Female(F)
X_7	Marital Status : Single(S), Monogamous(M), Polygamous(P), Widower(W), Divorced(D), Common-Law(C)
X_{36}	General Economic Situation of the Community : Worse Now(W), Bad Now(Ba), Unchanged(U), Better Now(Be), More Better Now(M)
X_{37}	How many people most contribute to the Household Expenses : 0, 1, 2, . . .
Housing	
X_8	Housing Tenure Status : Homeowner(H), Leaseholder(L), occupier rent free(O)
X_9	Number of Separate Rooms : 0, 1, 2, . . .
X_{10}	Has a Untreated Net : Yes(Y), No(N)
X_{11}	Has a Treated Net : Yes(Y), No(N)
X_{12}	Has an Iron : Yes(Y), No(N)
X_{13}	Has an Improved Stove : Yes(Y), No(N)
X_{14}	Has a Fridge/Freezer : Yes(Y), No(N)
X_{15}	Has a Television : Yes(Y), No(N)
X_{16}	Has a Bed/Mattress : Yes(Y), No(N)
X_{17}	Has a Modern Cooker : Yes(Y), No(N)
X_{18}	Has a Computer : Yes(Y), No(N)
X_{19}	Has Electricity : Yes(Y), No(N)
X_{20}	Is There a Room Equipped for Cooking ? : Yes(Y), No(N)
X_{21}	Has a Landline : Yes(Y), No(N)
X_{22}	Has a Mobile Phone : Yes(Y), No(N)
X_{23}	Area of Residence : Developed Area(D), Undeveloped Area(U)
X_{24}	Type of Housing : Apartment Building(A), Villa(V), Single Individual House(S), Multiple Occupancy Building(M), Traditional House(T), Other(O)
X_{25}	Soil Materials of the Main Building : Tile(T), Cement(C), Sand/Clay(S), Other(O)
X_{26}	Wall Materials of the House : Mud Brick/Earth(E), Stones(S), Bricks(B), Cement/Concrete(C), Wood/Bamboo(W), Metal Sheet(M), Wall Straw(WS), Other(O)
X_{27}	Materials of the Roof of the House : Earth(E), Straw(S), Bricks(B), Metal Sheet(M), Cement/Concrete(C), Tiles(T)
X_{28}	Main Energy Sources for Cooking : Firewood with Improved Stove(FI), Firewood with Single Stove(FS), Charbon with Improved Stove(CI), Charbon with Single Stove(CS), Kerosene/Oil(K), Gas(G), Electricity(E), Harvest Residue/Sawdust(H), Animal Waste(A), Other(O)
X_{29}	Main Energy Sources for Lighting : Kerosene/Oil(K), Gas(G), Electricity(E), Solar Energy(S), Generator(Ge), Battery Torch(B), Lamp Loadable/Batteries(L), Candle(C), Other(O)
Mobility	
X_{30}	Owner a Bicycle : Yes(Y), No(N)
X_{31}	Owner of a Motorcycle : Yes(Y), No(N)
X_{32}	Owner of a Car or a Lorry : Yes(Y), No(N)
X_{33}	Time Taken to Reach the Nearest Public Transport (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
Nutrition	
X_{34}	Has had Problems to Meet Food Needs : Never(N), Rarely(R), Sometimes(S), Often(O), Always(A)
X_{35}	General Economic Situation of the Household : Worse Now(W), Bad Now(Ba), Unchanged(U), Better Now(Be), More Better Now(M)
X_{38}	Time Taken to Reach the Market of Food Nearest (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
Water and Sanitation (WS)	
X_{39}	Main Source of Drinking Water : Barrage-River-Lake(B), Ordinary Wells(O), Closed Wells with Nozzles(C), Single Wells with Nozzles(S), Drilling(D), Public Fountain(P), Indoor Tap(I), Indoor Tap Shared (T), Other(O)
X_{40}	Type of Toilets used : Septic Tank(S), Traditional Latrines Rehabilitated(T), VIP Tank(V), Latrine or Toilet Flush(L), Ordinary Latrines(OL), In the Nature(I), Other(O)
X_{41}	Mode of Garbage Disposal : Dustbin(removal by service)(D), Public Garbage(P), Individual Garbage(I), Pits(Pi), Ferry(F), Road/Street(R), Other(O)
X_{42}	Mode of Sewage Disposal : In the Courtyard(C), In the Street/Road(S), In the Pit(P), Into Public Drains(D), In the Nature(I), Other(O)
X_{43}	Time Taken to Reach the Closest Source of Cooking Water (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
X_{44}	Time Taken to Reach the Closest Source of Drinking Water (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
Health	
X_{45}	Time Taken to Reach the Closest Health Service (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
Education	
X_{46}	Time Taken to Reach the Closest Primary School (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
X_{47}	Time Taken to Reach the Closest Secondary School (in minutes) : [0; 14] =Very Close(V), [15; 29] =Acceptably Close(A); [30; 44] = Close(C); [45; 59] =Far(F), 60+ =Far Away(FA)
X_{48}	Level of Education of Household Head : Has Never Been to School(N), Primary Not Completed(PN), Primary Completed(PC), Secondary Not Completed(SN), Secondary Completed(SC), Higher School(H), Adult Literacy(A)

TABLE A.8 – Description of criteria

Les techniques de clustering

B.1 Introduction

La classification non supervisée est abondamment utilisée dans des disciplines très diverses : biologie, psychologie, statistique et autres sciences sociales, dans le but d'identifier des groupes ou classes à partir d'un grand volume de données. Elle intervient notamment dans les applications du Data Mining (fouille des données), du Text Mining (classification du type de texte), du Information Retrieval (classification de documents), etc. Les techniques de classification sont utilisées dans de nombreux domaines tels que le domaine du marketing (segmentation de marché), le domaine politique (stratégie électorale en fonction du type d'électeurs), le domaine médical (le diagnostic médical), l'extraction d'information cachée ou enfouie dans de grand volume de données, etc. Les méthodes de classification font partie de la statistique exploratoire multidimensionnelle. La classification non supervisée est un processus consistant à organiser les objets en groupes d'éléments ou objets ou items tel que chaque groupe soit une collection d'objets "semblables" entre eux et "différents" des objets appartenant à d'autres groupes. Son objectif est d'explorer des données, de les réduire éventuellement, de découvrir des structures non visibles sur les données d'origine, de les valider en infirmant ou en confirmant une hypothèse soutenue avant de passer à l'étape de la modélisation. La classification organise les données par structuration abstraite sous-jacente au regroupement des individus ou à une hiérarchie de groupes. L'analyse de chaque groupe permet d'établir si les regroupements concordent avec la réalité ou si une nouvelle expérimentation est nécessaire pour conclure.

La partition fournie par une méthode de classification peut être constituée soit de classes réelles (données hétérogènes), soit de classes dites *instrumentales* (données homogènes), qui sont le résultat du découpage d'un ensemble continu. Ce dernier cas fait référence aux méthodes généralisant la construction d'histogrammes en statistique unidimensionnelle. Les méthodes de classification d'un ensemble d'individus peuvent se répartir en deux grands groupes : **classification par partition** (ou *partitionnement*) et **classification hiérarchique**. Les méthodes de partitionnement, qui exigent que le nombre de classes soit défini à l'avance, consistent généralement à diviser les données d'un ensemble en k classes disjointes. Les méthodes hiérarchiques produisent une séquence de partitions emboîtées sous forme

d'arbre communément appelé **dendrogramme** (*arbre hiérarchique*), visualisant ce système de classes organisées par inclusion.

Les deux grands types de méthodes peuvent être aussi divisées en trois groupes :

- (i) les méthodes fondées sur la notion de *distance*
- (ii) les méthodes fondées sur un *modèle*
- (iii) les méthodes fondées sur la notion de *densité*

Les méthodes (i) requièrent l'utilisation d'une distance ou d'un indice de dissimilarité pour effectuer le regroupement des individus homogènes dans la même classe. Les méthodes (ii) supposent que les données suivent une distribution statistique spécifique et que l'ensemble des données forme un mélange de distributions. L'approche neuronale, qui permet d'ajuster un modèle fortement non linéaire aux données, est un exemple d'approche de modèle. Les méthodes (iii) considèrent un cluster comme une région plus dense comparée à d'autres régions moins denses. La classification basée sur la densité est basée soit sur l'approche "*grille*", soit sur l'approche "*voisinage*". L'approche voisinage commence par fixer un rayon et définit ensuite le voisinage d'un objet comme étant le nombre minimal d'objet pour constituer un cluster autour de cet objet. Dans l'approche grille, l'espace des données est divisé en un ensemble fini de cellules multidimensionnelles ; la densité de chaque cellule est évaluée et les cellules denses et voisines l'une de l'autre sont ensuite regroupées dans un même cluster [Nakache et Confais, 2004].

Dans ce chapitre, nous commencerons par une présentation des généralités sur les approches de changement de base, les distances et indices de similarité communs aux différentes méthodes de classification. Nous terminerons ce chapitre par une revue de la littérature sur les méthodes de classification existantes.

B.2 Choix d'une mesure de ressemblance

B.2.1 Généralités

Les données soumises à une technique de classification se présentent le plus souvent sous la forme d'une matrice de réponses où les lignes désignent les individus (objets, instances, patterns, etc.) et les colonnes les variables (attributs, caractères, etc.). Par exemple, si les habitants d'un village doivent être classés, chaque ligne du tableau représenterait un individu et chaque colonne les réponses aux questions posées dans le formulaire d'enquête. L'information catégorielle ou extrinsèque tels que l'âge, le sexe, la religion ou la couleur de cheveux, est normalement utilisée pour interpréter les résultats d'une classification ; elle ne fait généralement pas partie du tableau des réponses. Les p variables sont habituellement

représentées comme un ensemble d'axes orthogonaux et les n individus¹ sont alors des vecteurs à p composantes appartenant à un espace appelé *espace des patterns*. Un cluster peut être visualisé comme une collection d'individus qui sont plus près les uns des autres. Un algorithme de classification consistera alors à identifier de tels regroupements "naturels" dans un espace des individus.

Les méthodes de classification nécessitent qu'un *indice de proximité* (ou de ressemblance ou d'affinité) soit établi entre les paires d'individus. Cet indice peut être calculé à partir d'un tableau de réponses ou à partir des données brutes. Par exemple, plusieurs clients peuvent être appelés à ranger leurs préférences pour les marques de savon. La proximité entre deux marques de savon peut être calculée en faisant la moyenne sur les individus. On peut aussi demander à chaque individu de fournir directement les indices de proximité – par jugement de la similarité – entre les marques de savon sur une échelle de 1 à 10.

Soit \mathbf{x}_i , le vecteur représentant l'évaluation de l'individu i sur m attributs. Une matrice de proximité $[\delta(\mathbf{x}_i, \mathbf{x}_k)]$ contient les indices de proximité $\delta(\mathbf{x}_i, \mathbf{x}_k)$; où la paire $(\mathbf{x}_i, \mathbf{x}_k)$ représente le i ème individu de la i ème ligne et le k ème individu de la k ème colonne de la matrice de proximité avec $\mathbf{x}_i, \mathbf{x}_k \in \Omega$, où Ω est l'ensemble des individus. Les éléments de la diagonale principale d'une matrice de proximité ne sont pas pris en compte pendant l'analyse car le degré de proximité entre deux patterns identiques est le même. Nous supposons que toutes les matrices de proximité sont symétriques. Hubert et Levin [1976]; Gower et Zielman [1998] ont travaillé sur des matrices de proximité non-symétriques.

Définition B.2.1 Une distance d définie sur un ensemble d'individus Ω est une application $d : \Omega \times \Omega \rightarrow \mathbb{R}_+$ satisfaisant aux trois axiomes suivants :

- (i) Symétrie : $d(\mathbf{x}_i, \mathbf{x}_k) = d(\mathbf{x}_k, \mathbf{x}_i), \forall \mathbf{x}_i, \mathbf{x}_k \in \Omega$
- (ii) Positivité stricte : $d(\mathbf{x}_i, \mathbf{x}_k) \begin{cases} > 0, & \text{si } \mathbf{x}_i \neq \mathbf{x}_k \\ = 0, & \text{si } \mathbf{x}_i = \mathbf{x}_k \end{cases} \quad \forall \mathbf{x}_i, \mathbf{x}_k \in \Omega$
- (iii) Inégalité triangulaire : $d(\mathbf{x}_i, \mathbf{x}_k) \leq d(\mathbf{x}_i, \mathbf{x}_l) + d(\mathbf{x}_l, \mathbf{x}_k), \forall \mathbf{x}_i, \mathbf{x}_k, \mathbf{x}_l \in \Omega$

Définition B.2.2 Une similarité s définie sur un ensemble d'individus Ω est une application $s : \Omega \times \Omega \rightarrow \mathbb{R}_+$ satisfait, en plus des deux axiomes (i) et (ii) de la définition B.2.1, l'axiome suivant :

- (iv) $s(\mathbf{x}_i, \mathbf{x}_i) \geq \max_k s(\mathbf{x}_i, \mathbf{x}_k), \forall \mathbf{x}_i, \mathbf{x}_k \in \Omega.$

Définition B.2.3 Une dissimilarité δ est définie à partir de l'indice de similarité tel que :

- (v) $\delta(\mathbf{x}_i, \mathbf{x}_i) = 0, \forall \mathbf{x}_i \in \Omega$

1. Bien que certains utilisent le terme "objet" ou "pattern", nous avons trouvé plus "approprié" d'utiliser le mot "individu" plutôt que le mot "objet" car il est plus "moral" de prendre un objet pour un individu qu'un individu pour un objet.

$$(vi) \delta(\mathbf{x}_i, \mathbf{x}_k) = 1 - s(\mathbf{x}_i, \mathbf{x}_k), \forall \mathbf{x}_i, \mathbf{x}_k \in \Omega$$

Une dissimilarité qui vérifie seulement l'axiome (v) et (vi) est appelé *indice de dissimilarité*; s'il satisfait l'axiome (iii), on dit que c'est un *écart*.

Définition B.2.4 Une ultramétrie (indice ultramétrique) d_u est un indice de dissimilarité vérifiant l'inégalité suivante :

$$(vii) d_u(\mathbf{x}_i, \mathbf{x}_k) \leq \max(d_u(\mathbf{x}_i, \mathbf{x}_l), d_u(\mathbf{x}_l, \mathbf{x}_k)), \forall \mathbf{x}_i, \mathbf{x}_k, \mathbf{x}_l \in \Omega$$

Une ultramétrie satisfaisant (iii) est appelée *distance ultramétrique*. La proposition B.2.1 [voir Diday *et al.*, 1982, pour une démonstration] montre qu'il est toujours possible de transformer un indice de dissimilarité en une distance respectant le même ordre sur les couples.

Définition B.2.5 Une boule ultramétrique $\mathcal{B}(\mathbf{x}_i, \rho)$ de l'ensemble Ω muni d'une distance ultramétrique d_u , de centre \mathbf{x}_i et de rayon ρ est telle que :

$$\mathcal{B}(\mathbf{x}_i, \rho) = \{\mathbf{x}_k \in \Omega \text{ tel que } d_u \leq \rho\} \quad (\text{B.1})$$

On note $\mathcal{B}(\Omega)$, l'ensemble des boules ultramétriques de Ω :

$$\mathcal{B}(\Omega) = \{\mathcal{B}(\mathbf{x}_i, \rho) \text{ tel que } \rho \geq 0\} \quad (\text{B.2})$$

La boule $\mathcal{B}(\Omega)$ est dite "*ouverte*" si $d_u(\mathbf{x}_i, \mathbf{x}_k) < \rho$ et "*fermée*" si $d_u(\mathbf{x}_i, \mathbf{x}_k) \leq \rho$.

Proposition B.2.1 A tout indice de dissimilarité δ ne vérifiant pas l'inégalité triangulaire, on peut associer une distance d telle que :

$$\begin{aligned} \delta(\mathbf{x}_i, \mathbf{x}_j) < \delta(\mathbf{x}_l, \mathbf{x}_k) &\Leftrightarrow d(\mathbf{x}_i, \mathbf{x}_j) < d(\mathbf{x}_l, \mathbf{x}_k) \\ \delta(\mathbf{x}_i, \mathbf{x}_j) = \delta(\mathbf{x}_l, \mathbf{x}_k) &\Leftrightarrow d(\mathbf{x}_i, \mathbf{x}_j) = d(\mathbf{x}_l, \mathbf{x}_k) \quad \forall \mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_l, \mathbf{x}_k \in \Omega. \end{aligned}$$

B.2.2 Mesures de ressemblance entre individus

On appelle *mesure de ressemblance*, toute application à valeurs numériques qui permet d'exprimer numériquement un lien existant entre les individus ou entre les variables ; Les deux grands types de méthodes de classification utilisent différentes distances et indices de similarité [Jain et Dubes, 1988].

B.2.2.1 Cas des tableaux où toutes les variables sont continues

On associe au i ème individu de Ω un vecteur $\mathbf{x}_i = \begin{pmatrix} x_i^1 \\ x_i^2 \\ \vdots \\ x_i^p \end{pmatrix}$ et $X_j = \begin{pmatrix} x_1^j \\ x_2^j \\ \vdots \\ x_n^j \end{pmatrix}$

Les distances euclidiennes (ou quadratiques) peuvent se mettre sous la forme :

$$d^2(\mathbf{x}_i, \mathbf{x}_k) = {}^t(\mathbf{x}_i - \mathbf{x}_k) \mathcal{Q}(\mathbf{x}_i - \mathbf{x}_k) \quad (\text{B.3})$$

où \mathcal{Q} est une matrice symétrique définie positive. Citons-en quelques unes.

- La distance de Mahalanobis

La distance de Mahalanobis est donnée par l'équation B.4 suivante :

$$d^2(\mathbf{x}_i, \mathbf{x}_k) = {}^t(\mathbf{x}_i - \mathbf{x}_k)\Sigma^{-1}(\mathbf{x}_i - \mathbf{x}_k) \quad (\text{B.4})$$

où Σ est la matrice de variance-covariance² des individus.

- La distance du Khi-deux

Cette distance est particulièrement bien adaptée aux tableaux de fréquence. Elle est à la base de l'analyse factorielle des correspondances. Sa formulation est donnée par l'équation (B.5) suivante :

$$d^2(\mathbf{x}_i, \mathbf{x}_k) = \sum_{j=1}^p \frac{1}{X_j^+} \left(\frac{x_i^j}{\mathbf{x}_i^+} - \frac{x_k^j}{\mathbf{x}_k^+} \right)^2 \quad \text{où } X_j^+ = \sum_{i=1}^n x_i^j \text{ et } \mathbf{x}_i^+ = \sum_{j=1}^p x_i^j \quad (\text{B.5})$$

- La distance de Minkowsky

Elle est donnée par l'équation (B.6) suivante :

$$d(\mathbf{x}_i, \mathbf{x}_k) = \left[\sum_{j=1}^p |x_i^j - x_k^j|^\alpha \right]^{1/\alpha} \quad \text{où } \alpha \geq 1 \quad (\text{B.6})$$

La distance de Minkowsky est dite *semi-euclidienne* car pour certaines valeurs de α , on ne peut pas mettre cette distance sous forme quadratique. Par exemple, les trois distances ci-dessous sont non quadratiques :

– la distance de Manhattan (ou “city block” distance) :

$$d(\mathbf{x}_i, \mathbf{x}_k) = \sum_{j=1}^p \lambda_j |x_i^j - x_k^j| \quad (\text{B.7})$$

où λ_j est le poids de la j ème variable.

– la distance de Cambera :

$$d(\mathbf{x}_i, \mathbf{x}_k) = \sum_{j=1}^p \frac{|x_i^j - x_k^j|}{x_i^j + x_k^j} \quad (\text{B.8})$$

– la distance de Chebychev :

$$d(\mathbf{x}_i, \mathbf{x}_k) = \max_j |x_i^j - x_k^j| \quad (\text{B.9})$$

Remarque B.2.1 Si les données sont numériques et issues de variables mesurées dans une échelle non linéaire (par exemple e^x), on peut procéder à la transformation logarithmique et calculer les distances à partir de ces données numériques transformées.

2. $\Sigma = [\text{cov}(X_j, X_{j'})]_{j,j' \in \{1;2;\dots;p\}} = \text{var}(X) = E[(X - E(X))(X - E(X))^t]$ où E est l'espérance mathématique et $X = (X_1, X_2, \dots, X_p)$.

- Angle entre deux individus \mathbf{x}_i et \mathbf{x}_k

Cette distance est donnée par l'équation suivante :

$$d(\mathbf{x}_i, \mathbf{x}_k) = \arccos \left(\frac{{}^t \mathbf{x}_i \mathbf{x}_k}{\|\mathbf{x}_i\| \cdot \|\mathbf{x}_k\|} \right) \quad (\text{B.10})$$

où $\|\mathbf{x}_i\|$ désigne la norme du vecteur \mathbf{x}_i et ${}^t \mathbf{x}_i$ le vecteur transposé. Cette mesure de proximité est utilisée dans des applications où les individus sont de norme 1 (*Manhattan norm*³ en anglais), de telle sorte que la mesure de proximité est un angle entre les deux individus \mathbf{x}_i et \mathbf{x}_k .

Remarque B.2.2 Très souvent, la similarité entre deux individus \mathbf{x}_i et \mathbf{x}_k est donnée par la formule : $s(\mathbf{x}_i, \mathbf{x}_k) = \frac{1}{1 + d(\mathbf{x}_i, \mathbf{x}_k)}$ où $d(\mathbf{x}_i, \mathbf{x}_k)$ est la distance choisie entre deux points \mathbf{x}_i et \mathbf{x}_k . D'autres indices de similarité sont donnés par les équations (B.11), (B.12) et (B.13) :

$$s_{\cos}(\mathbf{x}_i, \mathbf{x}_k) = \frac{{}^t \mathbf{x}_i \mathbf{x}_k}{\|\mathbf{x}_i\| \cdot \|\mathbf{x}_k\|} \quad (\text{B.11})$$

$$s_{\text{Dice}}(\mathbf{x}_i, \mathbf{x}_k) = 2 \cdot \frac{{}^t \mathbf{x}_i \mathbf{x}_k}{\|\mathbf{x}_i\|^2 + \|\mathbf{x}_k\|^2} \quad (\text{B.12})$$

$$s_{\text{exp}}(\mathbf{x}_i, \mathbf{x}_k) = \exp \left(- \|\mathbf{x}_i - \mathbf{x}_k\|^\alpha \right) \quad (\text{B.13})$$

où α est une valeur quelconque.

B.2.2.2 Cas des tableaux binaires

Pour définir une mesure de ressemblance lorsque \mathbf{x}_i et \mathbf{x}_k sont deux vecteurs binaires, il est nécessaire d'introduire les quatres quantités suivantes :

- p_{11} le nombre de fois où $x_i^j = x_k^j = 1$
- p_{12} le nombre de fois où $x_i^j = 0$ et $x_k^j = 1$
- p_{21} le nombre de fois où $x_i^j = 1$ et $x_k^j = 0$
- p_{22} le nombre de fois où $x_i^j = x_k^j = 0$

Les mesures de ressemblance suivantes ont souvent été utilisées :

Remarque B.2.3

3. $\|\mathbf{x}_i\|_1 = |x_i^1| + |x_i^2| + \dots + |x_i^p|$, et induit la distance de déplacement à angle droit sur un damier.

$\delta_1(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11}}{p_{11} + p_{12} + p_{21} + p_{22}}$	(Russell et Rao)
$\delta_2(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11}}{p_{11} + p_{12} + p_{21}}$	(Jaccard, Needham)
$\delta_3(\mathbf{x}_i, \mathbf{x}_k) = \frac{2p_{11}}{2p_{11} + p_{12} + p_{21}}$	(Dice)
$\delta_4(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11}}{p_{11} + 2(p_{12} + p_{21})}$	(Sokal et Sneath)
$\delta_5(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11} + p_{22}}{p_{11} + p_{12} + p_{21} + p_{22}}$	(Sokal et Michener)
$\delta_6(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11}}{p_{12} + p_{21}}$	(Kulzinsky)
$\delta_7(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11} + p_{22}}{p_{11} + 2(p_{12} + p_{21}) + p_{22}}$	(Rogers et Tanimoto)
$\delta_8(\mathbf{x}_i, \mathbf{x}_k) = \frac{p_{11}p_{22} - p_{12}p_{21}}{p_{11}p_{22} + p_{12}p_{21}}$	(Yule)
$\delta_9(\mathbf{x}_i, \mathbf{x}_k) = \frac{ p_{11}p_{22} - p_{12}p_{21} }{[(p_{11} + p_{12})(p_{21} + p_{22})(p_{11} + p_{21})(p_{12} + p_{22})]^{1/2}}$	(Corrélation)

TABLE B.1 – Quelques mesures de ressemblance pour les tableaux binaires

1. Les mesures de ressemblance δ_5 , δ_7 , $\delta'_8 = \delta_8 + 1$ et $\delta'_9 = \delta_9 + 1$ sont des indices de similarité; les autres mesures de ressemblance δ_j ne satisfont pas généralement à la condition $\delta_j(\mathbf{x}_i, \mathbf{x}_k) = \delta_j(\mathbf{x}_k, \mathbf{x}_i) \forall (\mathbf{x}_k, \mathbf{x}_i) \in \Omega^2$.
2. Les coefficients p_{11} , p_{12} , p_{21} , p_{22} peuvent se formuler de la façon suivante :

$${}^t\mathbf{x}_i \cdot \mathbf{x}_k = p_{11} \quad (\text{B.14})$$

$$\|\mathbf{x}_i - \mathbf{x}_k\|^2 = \sum_{j=1}^p \|x_i^j - x_k^j\|^2 = p_{12} + p_{21} \quad (\text{B.15})$$

$$d = n - {}^t\mathbf{x}_i \cdot \mathbf{x}_k - \|\mathbf{x}_i - \mathbf{x}_k\|^2 \quad (\text{B.16})$$

B.2.2.3 Cas de tableaux où toutes les variables sont qualitatives nominales

Pour calculer la ressemblance entre les individus, on peut transformer des données qualitatives en données binaires par codage disjonctif complet [cf. [Diday et al., 1982](#)]. Ensuite, appliquer les mesures de ressemblance indiquées pour les variables continues ou binaires au nouveau tableau obtenu. On peut aisément remarquer, du point de vue informatique, que le problème de la taille du tableau binaire obtenu se pose. Par exemple, pour un tableau de 500 individus contenant 40 variables nominales à 10 modalités, on passe d'un tableau initial de 20 000 cases (soit 500×40) à un tableau de 200 000 cases.

Pour éviter ce problème, [Diday et al. \[1982\]](#) proposent d'utiliser des distances sur le tableau binaire sans toutefois le charger en mémoire. Pour cela, on utilise uniquement les données contenues dans le tableau initial où seules les modalités sont

représentées. Les propriétés peuvent alors être utilisées pour calculer les distances entre les paires d'individus.

Proposition B.2.2 *La distance euclidienne classique utilisée sur le tableau binaire peut s'exprimer comme suit :*

$$d^2(\mathbf{x}_i, \mathbf{x}_k) = 2 \sum_{j=1}^p \delta_{ik}^j \quad \text{où } \delta_{ik}^j = \begin{cases} 1, & \text{si } x_i^j \neq x_k^j; \\ 0, & x_i^j = x_k^j. \end{cases} \quad (\text{B.17})$$

Proposition B.2.3 *La distance de Khi-deux χ^2 utilisée sur le tableau binaire peut aussi s'exprimer comme suite :*

$$d^2(\mathbf{x}_i, \mathbf{x}_k) = \frac{1}{p^2} \sum_{j=1}^p \left(\frac{1}{\#(x_i^j)} + \frac{1}{\#(x_k^j)} \right) \delta_{ik}^j \quad \text{où } \delta_{ik}^j = \begin{cases} 1, & \text{si } x_i^j \neq x_k^j; \\ 0, & x_i^j = x_k^j. \end{cases} \quad (\text{B.18})$$

et $\#(x_i^j)$ est le nombre de fois que la modalité x_i^j est atteinte par la jème variable.

B.2.2.4 Cas de tableaux où toutes les variables sont qualitatives ordinales

Dans ce cas, on remplace les valeurs x_i^j par leurs rangs r_{ij} ($r_{ij} = \{1; 2; \dots; \kappa_j\}$), où κ_j est le nombre de valeurs distinctes de la variable X_j . Puis, on utilise la transformation $\mathcal{Z}_{ij} = \frac{r_{ij} - 1}{\kappa_j - 1}$ qui fournit une variation de \mathcal{Z}_{ij} entre 0 et 1. Dès lors, on calcule les distances entre individus à partir des \mathcal{Z}_{ij} en utilisant les mesures de ressemblance qui ont été indiquées dans le cas où toutes les variables sont continues.

B.2.2.5 Cas de tableaux où toutes les variables sont mixtes

Pour calculer la distance, deux méthodes sont généralement utilisées. La première méthode recommande de transformer le tableau de données initiales en un tableau disjonctif complet et utiliser ensuite la distance de Khi-deux pour mesurer la ressemblance entre les individus. La deuxième utilise dans le calcul des distances entre individus et une formule (équation) pondérée pour combiner les effets des différentes variables :

$$d(\mathbf{x}_i, \mathbf{x}_k) = \frac{\sum_{j=1}^p \delta_{ik}^{(j)} d_{ik}^{(j)}}{\sum_{j=1}^p \delta_{ik}^{(j)}} \quad (\text{B.19})$$

où :

(a) $\delta_{ik}^{(j)} = \begin{cases} 0, & \text{si soit } x_i^j \text{ ou } x_k^j \text{ est une donnée manquante;} \\ 1, & \text{sinon.} \end{cases}$

(b) $d_{ik}^{(j)}$ est le type de distance qui dépend de la nature des variables.

- si la j ème variable est binaire ou nominale alors : $d_{ik}^{(j)} = \begin{cases} 0, & \text{si } x_i^j = x_k^j ; \\ 1, & \text{sinon.} \end{cases}$
- si la j ème variable est continue : $d_{ik}^{(j)} = \frac{|x_i^j - x_k^j|}{\max(X^j) - \min(X^j)}$
- si la variable est ordinale : $d_{ik}^{(j)}$ est la distance normalisée calculée à partir des scores $r_{ik}^{(j)}$.

B.2.3 Mesures de ressemblance entre variables

Dans le cas des données continues ou binaires, les mesures de ressemblance entre individus et les mesures de ressemblance entre variables sont du même type. Ainsi, nous distinguons ici deux cas :

B.2.3.1 Cas des variables nominales

A partir de deux variables nominales v_p et v_q (avec $(p, q) \in \mathcal{J}^2$), on définit le tableau de contingence (voir Table B.2) entre ces deux variables. Dans ce tableau, n_{ij} est le nombre d'individus ayant simultanément répondu à la i ème modalité de la variable v_p et la j ème modalité de la variable v_q . D'autre part, $n_{\bullet\bullet} = \sum_{i=1}^l \sum_{j=1}^r n_{ij}$; $n_{\bullet j} = \sum_{i=1}^l n_{ij}$; $n_{i\bullet} = \sum_{j=1}^r n_{ij}$. Ce tableau (voir Table B.2) nous

$v_p \backslash v_q$	1	2	...	j	...	r	Total
1	n_{11}	n_{12}	...	n_{1j}	...	n_{1r}	$n_{1\bullet}$
2	n_{21}	n_{22}	...	n_{2j}	...	n_{2r}	$n_{2\bullet}$
...
i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{ir}	$n_{i\bullet}$
...
l	n_{l1}	n_{l2}	...	n_{lj}	...	n_{lr}	$n_{l\bullet}$
Total	$n_{\bullet 1}$	$n_{\bullet 2}$...	$n_{\bullet j}$...	$n_{\bullet r}$	$n_{\bullet\bullet}$

TABLE B.2 – Tableau de contingence

permet de calculer de nombreuses mesures de ressemblance, dont la plus classique est le χ^2 de contingence :

$$\delta_{10}(v_p, v_q) = \chi^2 = n_{\bullet\bullet} \sum_{i=1}^l \sum_{j=1}^r \left(n_{ij} - \frac{n_{i\bullet} n_{\bullet j}}{n_{\bullet\bullet}} \right)^2 \frac{1}{n_{i\bullet} n_{\bullet j}} \quad (\text{B.20})$$

Les deux variables v_p et v_q sont d'autant plus proches (c'est-à-dire, leur ressemblance est d'autant plus grande) que $\delta_{10}(v_p, v_q)$ est grand. Pour calculer une mesure de ressemblance liée au χ^2 et variant entre 0 et 1, il suffit d'utiliser le T de Tschuprow :

$$\delta_{11}(v_p, v_q) = T = \left[\frac{\theta^2}{[(l-1)(r-1)]^{1/2}} \right]^{1/2} \quad \text{où } \theta^2 = \frac{\chi^2}{n_{\bullet\bullet}}. \quad (\text{B.21})$$

Les trois mesures de ressemblance suivantes sont également basées sur le tableau de contingence (Table B.2) :

$$\delta_{12}(v_p, v_q) = C = \left[\frac{\chi^2}{n_{\bullet\bullet} \cdot \min[(l-1), (r-1)]} \right]^{1/2} = \left[\frac{\theta^2}{\min[(l-1), (r-1)]} \right]^{1/2} \quad (\text{B.22})$$

$$\delta_{13}(v_p, v_q) = \sum_{i=1}^l n_{i\bullet} \sum_{j=1}^r \left| \frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n_{\bullet\bullet}} \right| \quad (\text{B.23})$$

$$\delta_{14}(v_p, v_q) = \frac{1}{n_{\bullet\bullet}} \sum_{i=1}^l \max_j n_{ij} \quad (\text{B.24})$$

La mesure de ressemblance donnée par l'équation (B.22) est appelé *coefficient C de Cramer*. La distance basée sur la théorie de l'information peut aussi être calculée en utilisant le tableau de contingence. Elle s'obtient par la formule de l'équation (B.25) suivante :

$$\delta_I(v_p, v_q) = I(v_p) + I(v_q) - I(v_p \cap v_q) \quad (\text{B.25})$$

où

$$\begin{aligned} I(v_p) &= - \sum_{i=1}^l \frac{n_{i\bullet}}{n_{\bullet\bullet}} \log_2 \frac{n_{i\bullet}}{n_{\bullet\bullet}} \\ I(v_q) &= - \sum_{j=1}^r \frac{n_{\bullet j}}{n_{\bullet\bullet}} \log_2 \frac{n_{\bullet j}}{n_{\bullet\bullet}} \\ I(v_p \cap v_q) &= \sum_{i=1}^l \sum_{j=1}^r \frac{n_{ij}}{n_{\bullet\bullet}} \log_2 \frac{n_{ij}}{n_{\bullet\bullet}} \end{aligned}$$

Le coefficient proposé par Bacelar-Nicolau [1985, 1990] permet de calculer l'affinité entre deux variables. Ce coefficient est donné par l'équation (B.26) suivante :

$$\beta_{v_p v_q} = \sum_{i=1}^n \frac{n_{ip} n_{iq}}{n_{\bullet p} n_{\bullet q}} \quad (\text{B.26})$$

où $n_{\bullet p} = \sum_{i=1}^n n_{ip}$ est la marge de la j ème colonne. La distance de Hellinger, définie par l'équation (B.27), est liée au coefficient d'affinité par la relation donnée

par l'équation (B.28).

$$d_H^2(v_p, v_q) = \sum_{i=1}^n \left[\sqrt{\frac{n_{ip}}{n_{\bullet p}}} - \sqrt{\frac{n_{iq}}{n_{\bullet q}}} \right]^2 \quad (\text{B.27})$$

$$d_H^2 = 2[1 - \beta_{v_p v_q}] \quad (\text{B.28})$$

B.2.3.2 Cas des variables qualitatives ordinales

Pour comparer les variables qualitatives ordinales, on utilise une mesure de ressemblance entre ordres tels que : les coefficients de corrélation des rangs de Kendall, de Spearman ou de Guttman [cf. [Diday et al., 1982](#)].

- Corrélation des rangs de Kendall

Considérons deux variables v_p et v_q , l'ensemble des couples d'individus (i, k) avec $i < k$ telles que : $x_i^p \mathcal{R}_{v_p} x_k^p$ et $x_i^q \mathcal{R}_{v_q} x_k^q$. \mathcal{R}_{v_p} et \mathcal{R}_{v_q} sont deux relations pouvant être remplacées par les signes de l'ensemble $\{<, >, =\}$.

Si $\mathcal{R}_{v_p} \equiv \mathcal{R}_{v_q}$, on dit qu'il y a *concordance*.

Si $\mathcal{R}_{v_p} \not\equiv \mathcal{R}_{v_q}$, on dit qu'il y a *discordance*.

Soit ω le nombre de concordances et $\bar{\omega}$ le nombre de discordances. *Le coefficient de corrélation des rangs de Kendall* (ou τ de Kendall (1938)) s'écrit dans le cas où les applications v_p et v_q induisent des ordres totaux (c'est-à-dire v_p et v_q sont injectives).

$$\delta_{15}(v_p, v_q) = \tau_{v_p v_q} = \frac{2}{n(n-1)}(\omega - \bar{\omega}) \quad \text{avec } n = |\Omega|. \quad (\text{B.29})$$

Lorsque $\delta_{15}(v_p, v_q) = 1$, on parle de *concordance complète* et lorsque $\delta_{15}(v_p, v_q) = -1$, on parle de *discordance complète*.

Ce coefficient peut aussi s'exprimer en terme de corrélation de la façon suivante : soit l'application $\psi^j : \Omega \times \Omega \rightarrow \{-1; 0; 1\}$ telle que pour tout $(x_i, x_k) \in \Omega \times \Omega$ et $j \in \mathcal{J}$:

$$\begin{aligned} \psi^j(i, k) &= -1 \quad \text{où } x_i^j < x_k^j \\ \psi^j(i, k) &= 0 \quad \text{où } x_i^j = x_k^j \\ \psi^j(i, k) &= 1 \quad \text{où } x_i^j > x_k^j \end{aligned}$$

On obtient alors :

$$\delta_{15}(v_p, v_q) = \text{cor}(\psi^p, \psi^q) \quad (\text{B.30})$$

En guise de rappel, si X^p et X^q sont deux vecteurs de \mathbb{R}^n , $\text{cor}(X^p, X^q) = \frac{\text{cov}(v_p, v_q)}{\sigma(v_p)\sigma(v_p)}$

et $\sigma(v_p) = \sqrt{\text{var}(v_p)} = \left[\sum_{i=1}^n \lambda_i (x_i^p - \bar{v}_p)^2 \right]^{1/2}$ et $\text{cov}(v_p, v_q)$ est la *covariance* du

couple des vecteurs colonnes v_p et v_q donnée par l'équation B.31 suivante :

$$\text{cov}(v_p, v_q) = \sum_{i=1}^n \lambda_i (x_i^p - \bar{v}_p)(x_i^q - \bar{v}_q) = \sum_{i=1}^n \lambda_i x_i^p x_i^q - \bar{v}_p \bar{v}_q \quad (\text{B.31})$$

$$\text{avec } \bar{v}_p = \sum_{i=1}^n \lambda_i x_i^p.$$

- Corrélation des rangs de Spearman

Un rang est attribué à chaque valeur de la variable v_p et on fait de même pour la variable v_q . On a ainsi un codage qui permet de définir deux nouvelles variables π_p et π_q . On peut associer à chaque valeur x_i^p (respectivement x_i^q) son rang noté π_{ip} (respectivement π_{iq}). On peut utiliser le rang moyen lorsqu'il y a d'ex-aequo.

Soit $\bar{\pi}_j = \frac{1}{n} \sum_{i=1}^n \pi_{ij}$, le coefficient de corrélation des rangs de Spearman (1904) s'écrit alors :

$$\Pi_S = \frac{\sum_{i=1}^n (\pi_i^p - \bar{\pi}_p)(\pi_i^q - \bar{\pi}_q)}{\left[\sum_{i=1}^n (\pi_i^p - \bar{\pi}_p)^2 \sum_{i=1}^n (\pi_i^q - \bar{\pi}_q)^2 \right]^{1/2}} \quad (\text{B.32})$$

Dans le cas où il n'y a pas d'ex-aequo, on en déduit la mesure de ressemblance donnée par l'équation B.33 suivante :

$$\delta_{16}(v_p, v_q) = \Pi_S = 1 - 6 \sum_{i=1}^n \frac{(\pi_i^p - \pi_{iq})^2}{n(n^2 - 1)} \quad (\text{B.33})$$

- Le coefficient de corrélation des rangs de Guttman (1974)

Pour calculer ce coefficient, on commence par définir α_{ik} et β_{ik} par :

$$\alpha_{ik} = \begin{cases} -1, & \text{si } x_i^p < x_k^p; \\ 0, & \text{si } x_i^p = x_k^p; \\ 1, & \text{si } x_i^p > x_k^p; \end{cases} \quad \beta_{ik} = \begin{cases} -1, & \text{si } x_i^q < x_k^q; \\ 0, & \text{si } x_i^q = x_k^q; \\ 1, & \text{si } x_i^q > x_k^q; \end{cases}$$

Considérons ensuite :

$$\theta_{ik}^{(1)} = |\alpha_{ik} \beta_{ik}|,$$

$$\theta_{ik}^{(2)} = |\alpha_{ik}|,$$

$$\theta_{ik}^{(3)} = |\beta_{ik}|,$$

$$\theta_{ik}^{(4)} = |\alpha_{ik}| + |\beta_{ik}| - |\alpha_{ik} \beta_{ik}|,$$

Soit $\omega_{ik}^{(m)}$ un nombre non négatif non nul si $\theta_{ik}^{(m)}$ est non nul. Le coefficient de monotonie s'écrit alors :

$$\mu^{(m)} = \frac{\sum_{i=1}^n \sum_{k=1}^n \omega_{ik}^{(m)} \alpha_{ik} \beta_{ik}}{\sum_{i=1}^n \sum_{k=1}^n \omega_{ik}^{(m)} \theta_{ik}^{(m)}} \quad (\text{B.34})$$

On démontre aisément que $-1 < \mu^{(m)} < 1$. Un coefficient de monotonie $\mu^{(m)}$ est dit “faible” (respectivement “forte”) lorsque $m = 1$ (respectivement, $m = 4$). Suivant le choix de m et $\omega_{ik}^{(m)}$, on pourra définir différentes familles de coefficients et retrouver comme cas particulier celui de Kendall et de Spearman et en déduire de nouvelles mesures de ressemblance entre variables qualitatives.

B.3 Les approches de clustering

La pratique de la classification des individus selon leurs proximités est la base de beaucoup de sciences. L’organisation des données en clusters homogènes est l’une des stratégies fondamentales d’apprentissage et de compréhension des données. La classification non supervisée (ou clustering, en anglais) est l’étude formelle des algorithmes et des méthodes de classification des individus. Elle a pour but de regrouper des individus en clusters homogènes en fonction de l’analyse des caractéristiques qui décrivent les individus. Un individu est décrit soit par un ensemble de mesures (ensemble des évaluations sur chaque variable), soit par des relations entre l’individu et les autres individus (indices de proximité). L’objectif de la classification non supervisée est simplement de trouver une organisation “convenable” et “valide” des individus sans chercher à établir des règles de séparation des individus en catégories.

Les clusters (ou classes ou groupes) homogènes sont des regroupements des individus qui se ressemblent et séparés de ceux qui sont dissemblables. [Everitt \[1980\]](#) propose trois principales définitions du mot “cluster” :

1. Un cluster est un ensemble d’entités semblables et différents des entités des autres clusters.
2. Un cluster est une agrégation de points dans l’*espace test* telle que la distance entre deux points quelconque du cluster soit inférieure à la distance entre tout point du cluster et tout point n’appartenant pas au cluster.
3. Les clusters peuvent être décrits comme des régions d’un espace multidimensionnel contenant une *densité de points relativement élevée*, séparées des autres régions par une zone contenant une *densité de points relativement faible*.

Les deux dernières définitions supposent que les individus devant être classés (de manière non supervisée) sont représentés dans un espace mesurable. Ainsi, nous reconnaissons un cluster lorsque nous l’observons dans un plan. Néanmoins, il n’est pas toujours aisé de dire comment les clusters ont été formés. En effet, bien qu’il soit facile de donner une définition fonctionnelle d’un cluster, il est très difficile de donner une définition opérationnelle d’un cluster. Ceci est due au fait que les individus peuvent être regroupés en clusters avec des objectifs différents à l’esprit.

Ainsi, une fois qu'on a défini la matrice de proximité, choisi la méthode de classification et fixé le nombre de clusters, il nous faut justifier l'existence de chaque cluster.

Dans cette thèse, nous nous intéressons particulièrement à deux grandes familles de méthodes de classification non supervisée, à savoir : les approches hiérarchiques et les approches par partitionnement [Jain et Dubes, 1988; Jain *et al.*, 1999; Nakache et Confais, 2004; Saporta, 2006; Tufféry, 2010]. Les approches hiérarchiques, qui produisent une séquence de clusters emboîtés d'hétérogénéité croissante de la plus fine à la plus grossière permettant de visualiser ce système de clusters organisés par inclusion. Les approches par partitionnement cherchent la meilleure partition en \mathcal{K} clusters disjoints d'individus, le nombre de classes (clusters ou groupes) \mathcal{K} étant fixé à priori. Le partitionnement se fait grâce à un processus itératif, fonction du nombre \mathcal{K} , qui affecte chaque individu au cluster le plus proche au sens d'une distance ou d'un indice de proximité en optimisant une certaine fonction objectif. Ces deux approches seront décrites plus en détails dans les sections B.3.1 et B.3.2 de ce chapitre qui se terminera par une présentation des méthodes de validation et de caractérisation des clusters existants.

B.3.1 Les approches hiérarchiques

Les approches hiérarchiques sont une famille de techniques qui génèrent des suites de clusters emboîtés les uns dans les autres pouvant être scindées en deux grands groupes : Classification Ascendante Hiérarchique (CAH) ou par agrégation et Classification Descendante Hiérarchique (CDH) ou par division. La CAH considère une matrice de proximité entre les individus et utilise un *algorithme agglomératif* qui suppose initialement que tous les individus sont des clusters ne contenant qu'un seul individu (singleton) ; Puis, à chaque étape du processus, elle forme initialement de petits clusters ne comprenant que des individus très semblables, ensuite, à partir de ceux-ci, elle fusionne les clusters les plus proches au sens d'une mesure d'agrégation, jusqu'à obtenir l'unique cluster contenant tous les individus. La CDH, quant à elle, inverse le processus de la CAH. La CDH repose sur un *algorithme divisif* muni d'un critère de division d'un sous-ensemble de variables, et qui procède par des divisions successives de l'ensemble des individus tout entier jusqu'à ce que soit atteint la partition triviale où chaque observation est un cluster. La division d'un cluster s'opère de façon à ce que la mesure d'agrégation entre les deux clusters descendants soit la plus grande possible, de manière à créer deux clusters bien séparés. Un autre mode de construction des classes a été proposé par Diday [1986], comme une généralisation des modèles hiérarchiques, est appelé la *classification pyramidale*. Les représentations pyramidales sont constituées, comme dans les dendrogrammes, des ensembles de parties appelées aussi *clusters* ou *paliers* de l'ensemble des individus à classer. Cependant, la représen-

tation pyramidale repose sur une structure graphique plus complexe des données du fait que deux clusters de la pyramide peuvent avoir une intersection non vide et ainsi certains individus à classer, peuvent appartenir à deux classes qui ne sont pas emboîtées l'une dans l'autre (classes empiétantes). Le dendrogramme obtenue dans ce cas est dit *hiérarchie de recouvrement* (ou *pyramide*).

B.3.1.1 La Classification Ascendante Hiérarchique

Un algorithme de classification ascendante hiérarchique consiste à rechercher à chaque étape les clusters les plus proches pour les fusionner, et l'étape la plus importante est celle liée au choix du critère d'agrégation entre deux clusters. En effet, il existe plusieurs méthodes d'agrégation possibles entre deux clusters et le choix d'une méthode dépend du critère d'agrégation que l'on se pose. Les algorithmes les plus classiques définissent la distance entre deux clusters à partir de la mesure de proximité entre les individus constituant chaque cluster.

Soient L_1 et L_2 deux sous-ensembles disjoints de Ω . Les critères usuels d'agrégation pour chaque couple de clusters peuvent ainsi être définis :

★ Le critère du saut minimal (single linkage)

Ce critère, dû à [Jardine et Sibson \[1971\]](#), considère que la distance entre deux clusters L_1 et L_2 est donnée par la distance minimum entre toutes les paires d'individus entre les deux clusters (paire composée d'un élément de chaque cluster).

$$d_{\min}(L_1, L_2) = \min\{d(\mathbf{x}_i, \mathbf{x}_k); (\mathbf{x}_i, \mathbf{x}_k) \in L_1 \times L_2\} \quad (\text{B.35})$$

Le "*single linkage*" a un intérêt en reconnaissance des formes puisqu'il permet de détecter des clusters *longs* et *sinueux*. Son point faible est qu'il est sensible à l'*effet de chaîne*⁴ et donc moins adapté pour la détection des clusters sphériques.

★ Le critère du diamètre (complete linkage)

Ce critère, dû à [Sørensen \[1957\]](#), considère que la distance entre deux clusters L_1 et L_2 est donnée par la distance maximum entre toutes les paires d'individus entre les deux clusters. Nous parlons aussi de *saut maximum*.

$$d_{\max}(L_1, L_2) = \max\{d(\mathbf{x}_i, \mathbf{x}_k); (\mathbf{x}_i, \mathbf{x}_k) \in L_1 \times L_2\} \quad (\text{B.36})$$

Le "*complete linkage*" est très sensible aux points aberrants. Contrairement au *single linkage*, ce critère ne possède pas de bonnes propriétés et n'est pas souvent utilisé en pratique [voir [Nakache et Confais, 2004](#); [Tufféry, 2010](#)].

4. Nous parlons d'effet de chaîne lorsque deux points très éloignés l'un de l'autre mais reliés par une suite de points très proches les uns des autres sont rassemblés dans le même clusters.

★ **Le critère de la moyenne (average linkage)**

Ce critère, dû à [Sokal et Michener \[1958\]](#), considère que la distance entre deux clusters L_1 et L_2 est donnée par la valeur moyenne des distances entre toutes les paires d'individus entre les deux clusters. On parle aussi de *saut moyen*.

$$d_{\text{moy}}(L_1, L_2) = \frac{1}{|L_1| \cdot |L_2|} \sum_{(\mathbf{x}_i, \mathbf{x}_k) \in L_1 \times L_2} d(\mathbf{x}_i, \mathbf{x}_k) \quad (\text{B.37})$$

Le “*average linkage*” est considéré comme un compromis entre le *single linkage* et le *complete linkage*.

★ **Le critère de Ward (centroid linkage)**

Ce critère, dû à [Ward \[1963\]](#), considère que la distance entre deux clusters L_1 et L_2 est donnée par la distance entre leur centre gravité. On parle aussi de *saut barycentrique* ou *perte d'inertie minimale*. Ce critère est lié à des calculs d'inertie à chaque étape de la procédure d'agrégation. Si on suppose que l'ensemble Ω à classer est un ensemble d'individus muni de masses d'un espace euclidien, on choisira alors de regrouper les deux clusters qui gênèrent le gain d'inertie intra-cluster minimum. Selon le théorème de Huygens, l'inertie total $\mathbf{T}(\Omega)$ d'un ensemble d'individus Ω est égale à la somme des inerties intra-classe \mathbf{W} (*Within Inertia* en anglais) et inter-classe \mathbf{B} (*Between Inertia* en anglais) de cet ensemble d'individus regroupés dans une partition \mathcal{P} (voir l'équation B.38).

$$\mathbf{T}(\Omega) = \mathbf{W}(\mathcal{P}) + \mathbf{B}(\mathcal{P}) \quad (\text{B.38})$$

où

$$\begin{aligned} \mathbf{T}(\Omega) &= \sum_{i=1}^n \lambda_i d^2(\mathbf{x}_i, g) \\ \mathbf{W}(\mathcal{P}) &= \sum_{k=1}^{\mathcal{K}} \mathbf{W}_k \quad \text{avec } \mathbf{W}_k = \sum_{x_i \in L_k} \lambda_i d^2(\mathbf{x}_i, g_k) \\ \mathbf{B}(\mathcal{P}) &= \sum_{k=1}^{\mathcal{K}} \mu_k d^2(g_k, g) \end{aligned}$$

avec \mathcal{K} représente le nombre de clusters L_l de masse μ_l ; λ_i est la masse de l'individu i ; $g = \sum_{i=1}^n \lambda_i \mathbf{x}_i$ et $g_k = \frac{1}{\mu_k} \sum_{\mathbf{x}_i \in L_k} \lambda_i \mathbf{x}_i$.

L'inertie intra-classe d'une partition mesure l'homogénéité de l'ensemble de clusters. Un cluster sera d'autant plus homogène que son inertie intra-classe sera faible (minimisation de \mathbf{W}). L'inertie inter-classe d'une partition mesure la séparation entre les classes qui la composent. Plus l'inertie inter-cluster d'une partition est grande, plus les clusters sont distinctement séparés (maximisation de \mathbf{B}) comme le montre la Figure B.1. On peut alors distinguer deux situations :

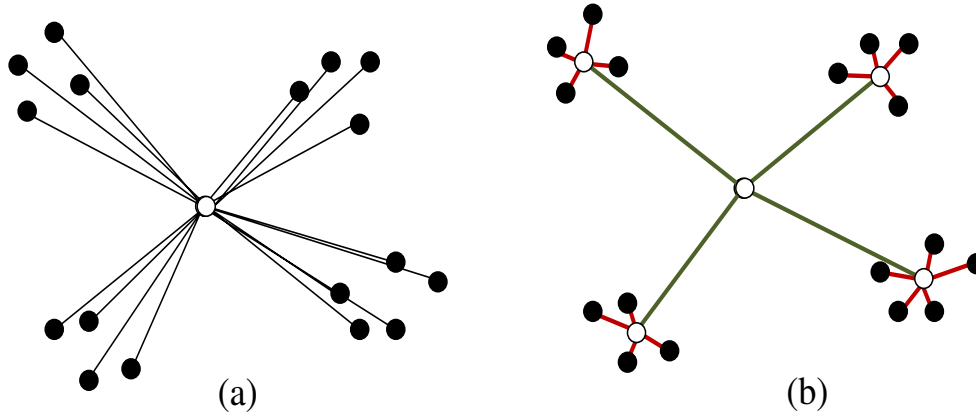


FIGURE B.1 – (a) Inertie Total = (b) Inertie Intercluster + Inertie Intra-cluster

- Si chaque cluster de la partition contient un seul individu, c'est-à-dire $\mathcal{K} = |\Omega| = n$, alors $\mathbf{W}(\mathcal{P}) = 0$ et $\mathbf{B}(\mathcal{P}) = \mathbf{T}(\Omega)$.
- Si la partition n'est composée que d'un seul cluster, c'est-à-dire $\mathcal{K} = 1$, alors $\mathbf{B}(\mathcal{P}) = 0$ et $\mathbf{W}(\mathcal{P}) = \mathbf{T}(\Omega)$.

On peut donc remarquer que l'inertie intra-cluster d'une partition décroît lorsque le nombre de clusters augmente dans la partition tandis que l'inertie inter-cluster décroît lorsque le nombre de clusters diminue dans la partition.

Ces critères classiques ont l'avantage d'être fondés sur un lien métrique qui les rend applicables à tout type de données dès lors que l'on est capable de construire une matrice de proximité entre les individus à classer. Ils présentent tous des inconvénients non négligeables dont le plus important est lié à la complexité algorithmique. En effet, la complexité de ces classifications n'est pas linéaire car pour passer de \mathcal{K} à $\mathcal{K} - 1$ clusters, nous devons calculer $\frac{\mathcal{K}(\mathcal{K} - 1)}{2}$ distances, réunir les deux classes les plus proches, puis recalculer les distances, avant de recommencer. Si n est le nombre d'individus à classer, la complexité des calculs des critères présentés est au moins en $\mathcal{O}(n^2)$ à l'exception de la méthode EML (Expected Maximum-Likelihood) qui est en $\mathcal{O}(n^3)$ [Nakache et Confais, 2004]. D'autre part, ces critères conduisent généralement à des partitions formées de clusters de forme convexe, de taille et densité sensiblement égales, sans tenir compte éventuellement des points aberrants.

De nouvelles méthodes de classification hiérarchique ont été développées, afin d'éviter les inconvénients mentionnés ci-dessus, et notamment, afin de fournir des partitions en classes de forme et taille arbitraires. Parmi celles-ci citons CURE⁵ (Clustering Using REpresentatives, [Guha *et al.*, 2001]), ROCK (Robust Clustering

5. Guha *et al.* [2001].

for Categorical Data, [Guha *et al.*, 2000]), BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies, [Zhang *et al.*, 1997]) et CHAMELEON (Hierarchical Clustering Algorithm Using Dynamic Modeling, [Karypis *et al.*, 1999]), que nous ne présentons pas dans cette thèse.

B.3.1.2 La classification descendante hiérarchique

Les méthodes de classification descendante hiérarchique sont itératives et procèdent à chaque itération au choix du segment de l'arbre hiérarchique à diviser, et au partitionnement de ce segment. Une méthode de classification 'divise' conduit à un arbre hiérarchique descendant dont les segments terminaux constituent une partition de l'ensemble des individus à classer. La différence entre les méthodes "divisives" réside dans les critères qu'elles utilisent pour choisir le segment à diviser ainsi dans la manière dont elles divisent le segment. Le choix de tels critères dépend généralement de la nature des variables caractérisant les individus à classer. On peut distinguer deux types d'approches divisives : les approches monothétiques et les approches conceptuelles :

★ Les approches monothétiques

Les approches monothétiques [Chavent, 1998; Chavent *et al.*, 2007] consistent à diviser l'ensemble des individus Ω en deux sous-segments (sous-clusters L_1 et L_2) suivant la réponse à la question de la forme "valeur de la variable $\leq \nu$?" : le sous-segment L_1 contient les individus pour lesquels la valeur de la variable est inférieure ou égale à ν et le sous-segment L_2 les individus pour lesquels la valeur de la variable est strictement supérieure à ν . L'algorithme divisif utilisé dépend du type de variables que l'on traite. Par exemple, dans le cas des variables continues, on calcule pour chaque division de la variable X_j , la variance intra-segment et on sélectionne la division suivant la variable X_j qui assure le minimum des différentes variances intra-segments $V_{\min}(X_j)$ pour tout j , c'est-à-dire, $\min\{V_{\min}(X_j); j = 1, \dots, p\}$. Pour les variables mixtes, on effectue un codage des données sous forme disjonctive complète afin d'utiliser une Analyse des Correspondances Multiples [Benzécri, 1992; Lebart *et al.*, 2006]. On arrête de diviser lorsque la taille de tous les sous-segments est inférieure ou égale à un nombre fixé.

L'inconvénient de cette approche divisive réside dans le fait qu'elle est très peu utile pour découvrir la structure en classes de l'échantillon analysé lorsque la structure de la classification dépend de combinaisons de variables et non des variables elles-mêmes. L'unique avantage des clusters monothétiques est qu'ils procurent la liste des caractères possédés (ou non possédés) en commun par tous les sujets appartenant à une même classe. Cela facilite énormément leur interprétation.

★ Les approches conceptuelles

L'approche conceptuelle a été introduite par Fisher [1987] comme une méthode divisive de classification automatique sur des données qualitatives. Elle est basée sur les deux fonctions PU (Partition Utility) et CU (Category Utility). La fonction PU (Partition Utility) mesure la qualité d'une partition en \mathcal{K} clusters $(L_1, L_2, \dots, L_{\mathcal{K}})$ de n individus sur lesquels sont relevées les mesures de p variables X_j (avec $j \in \mathcal{J}$ et $\mathcal{J} = 1, \dots, p$) qualitatives présentant \mathcal{M}_j modalités. La fonction CU (Category Utility) est définie comme l'accroissement du nombre attendu de modalités des variables correctement prédites par rapport au nombre de modalités prédites correctes sans connaissance d'une telle partition.

B.3.2 Les approches par partitionnement

Les approches par partitionnement fournissent une partition unique des éléments à classer et exigent que le nombre de clusters de la partition à construire soit fixé a priori. Ces approches consistent à constituer, à partir d'un ensemble de n individus $\mathbf{x}_i \in \Omega$ de masses respectives μ_i telles que $\sum_{i=1}^n \mu_i = 1$, des clusters d'individus (homogènes et distinctes) en nombre \mathcal{K} fixé. Elles sont basées sur une distance ou un indice de similarité entre clusters à classer. Les individus regroupés dans un même cluster sont très proches les uns des autres en termes de distance ou d'indice de similarité. Une partition optimale peut être obtenue si l'on parvient à énumérer de façon exhaustive toutes les partitions possibles. Ce qui est généralement impossible en pratique à cause d'un temps de calcul exponentiel. Toutefois, les approches par partitionnement basées sur l'optimisation itérative d'un certain critère permettent d'obtenir des clusters assez distincts en un temps de calcul polynomial. Il s'agit de différents algorithmes de réaffectation qui consistent à redistribuer itérativement les individus dans \mathcal{K} clusters. Ces algorithmes de réaffectation permettent d'améliorer progressivement la qualité des clusters et conduisent pour des données appropriés à des clusters de bonne qualité.

Les algorithmes de partitionnement par optimisation itérative sont divisés en méthodes des \mathcal{K} -means et méthodes des \mathcal{K} -medoids selon la définition des représentations des clusters. Ils tendent à construire des clusters de forme convexe sans tenir compte des *outliers* (valeurs aberrantes).

Les méthodes \mathcal{K} -means [Hartigan et Wong, 1979], très populaires dans les applications scientifiques et industrielles, permettent de scinder n individus multidimensionnels en \mathcal{K} partitions (avec $\mathcal{K} < n$) ou clusters ne se chevauchant pas. L'objectif de ces méthodes est de *minimiser* l'inertie intra-classe \mathbf{W} (encore appelée *fonction erreur carrée*) donnée par l'équation (B.38) où il existe \mathcal{K} clusters L_k et g_k est le point moyen (ou "centroïde" ou "prototype") de tous les points $\mathbf{x}_i \in L_k$. Le nombre \mathcal{K} de clusters est le paramètre d'entrée des méthodes \mathcal{K} -means.

La méthode des centres mobiles [Forgy, 1965] et la méthode des nuées dynamiques [Diday, 1971], présente dans plusieurs logiciels, sont apparentées à ces méthodes \mathcal{K} -means. La complexité des algorithmes \mathcal{K} -means est de l'ordre de $\mathcal{O}(iter.\mathcal{K}n)$. En plus du fait que les méthodes \mathcal{K} -means exigent que le nombre de clusters soit connu à l'avance, elles ne permettent pas de détecter des données bruitées ou la présence d'outliers et fournissent des clusters convexes. Le résultat de ces méthodes dépend fortement du tirage initial des points représentant les centroïdes et conduit ainsi à l'obtention d'une solution locale. Il est donc conseillé d'effectuer plusieurs tirages et de comparer les différentes partitions correspondantes obtenues, ce qui est préjudiciable au niveau de la complexité de ces méthodes.

Dans les méthodes \mathcal{K} -medoids, chaque cluster est représenté par un individu ou représentant ce cluster. Ce sont aussi des méthodes d'optimisation itérative combinant la réaffectation des points aux clusters avec l'interversion des représentants et des autres points. Cette stratégie est coûteuse du point de vue temps de calcul, donc à éviter pour des données de grande taille. Les méthodes PAM (Partition Around Medoids)[Kaufman et Rousseeuw, 2005], CLARA (Clustering LARge Application)[Kaufman et Rousseeuw, 2005], CLARANS (Clustering LARge Applications based upon RANdomized Search)[Ng et Han, 1994] et FINDIT (Fast and Intelligent Subspace Clustering using Dimension Voting)[Woo *et al.*, 2004] sont apparentées à ces méthodes \mathcal{K} -medoids. La complexité des algorithmes \mathcal{K} -medoids est de l'ordre de $\mathcal{O}(\mathcal{K}(n-\mathcal{K})^2)$ pour PAM, $\mathcal{O}(\mathcal{K}s^2 + \mathcal{K}(n-\mathcal{K}))$ pour CLARA (où s représente la taille de l'échantillon de données extraits de l'échantillon total) et $\mathcal{O}(\mathcal{K}n^2)$ pour CLARANS. La méthode FINDIT est très efficace pour des données hautement multidimensionnelles du fait qu'elle satisfait certaines propriétés importantes des qualités d'une bonne méthode de classification telles que la précision des résultats, la robustesse au bruit et à la densité, la complexité des calculs.

B.3.3 Les approches fondées sur la densité

Ces approches sont basées sur la densité des régions de l'espace et utilisent un critère de classification local. Un cluster est défini comme une région de l'espace expérimental où la densité dans le voisinage des individus est très importante. Les clusters sont des régions de l'espace des données de forte densité, séparées par des régions de faible densité ou sans densité. Les algorithmes fondés sur la densité regroupent les individus suivant des fonctions objectives de densité. La densité est le plus souvent définie comme le nombre d'individus dans un certain voisinage. Un cluster regroupe des individus tant que le nombre d'objets dans le voisinage ne dépasse pas une valeur fixée du paramètre. Le principe de ces approches est différent de celui des approches par partitionnement où le nombre \mathcal{K} de clusters à découvrir est fixé à l'avance et qui utilise la réaffectation des individus aux différents clusters.

Ces approches requièrent les concepts de *densité*, *noyau*, *point limite*, *accessibilité* et *connectivité* qui sont liés à la notion de plus proches voisins d'un point. Un cluster croît dans la direction où se trouve la densité. Les algorithmes correspondants sont donc aptes à découvrir des clusters de formes arbitraires et fournissent une protection naturelle contre les outliers. Ces approches peuvent être utilisées avec beaucoup de succès sur des données de grande taille.

L'objectif de ces approches est de reconnaître des clusters tels qu'à l'intérieur de chaque cluster on ait une densité de points beaucoup plus grande qu'à l'extérieur et la densité des points dans les zones de bruit soit plus petite que la densité des points dans n'importe quel cluster. L'idée clé est que pour chaque point d'un cluster, le voisinage autour de ce point dans un rayon donné ε contienne au moins un nombre minimum de points m_ε (c'est-à-dire la densité dans un voisinage doit dépasser un certain seuil). La forme du voisinage est déterminée par le choix d'une distance entre deux individus x_i et x_j de l'ensemble Ω . La méthode DBSCAN (Density-Based Spatial Clustering of Applications with Noise) [Ester *et al.*, 1996] et ses dérivées GDBSCAN [Sander *et al.*, 1998] et OPTICS (Ordering Points To identify the clustering Structure) [Ankerst *et al.*, 1999] sont apparentées aux méthodes fondées sur la densité. Ces méthodes tiennent compte des outliers et ont une complexité de l'ordre de $\mathcal{O}(n \log(n))$ pour OPTICS.

B.3.4 Les approches fondées sur un modèle

Ces approches fournissent de bonnes approximations des paramètres de modèles qui ajustent au mieux les données. Ces paramètres de modèles peuvent être de type partitionnement ou hiérarchique, dépendant de la structure ou du modèle sous-jacent et de la façon qu'ils ont de peaufiner le modèle pour identifier des clusters. Ce type d'approches est basé sur l'hypothèse forte que les données de chaque cluster suivent une distribution de probabilité spécifique et que l'ensemble des données forme un mélange de distributions. Cette hypothèse garantit la création des clusters particuliers tel que le modèle préconçu puisse être amélioré. Néanmoins, les algorithmes de ce type sont initialisés avec un nombre fixé de clusters et n'utilisent pas le même concept de densité bien qu'ils soient plus proches des algorithmes fondés sur la densité.

Les cartes auto-organisatrices (Self-Organizing Maps, SOM)[Kohonen, 1988; Kohonen *et al.*, 2001] dues à Kohonen, sont apparentées aux approches fondées sur un modèle. Elles sont construites par un algorithme très proche des algorithmes de centres mobiles. On représente les données par un nombre réduit de représentants (appelés neurones). L'approche d'apprentissage est compétitive : le "gagnant" est le représentant le plus proche d'un individu donné. La différence entre les SOM et la méthode des centres mobiles se situe au niveau de la préservation de la carte. Dans cette carte, les individus les plus semblables seront les plus proches sur la

carte : il y aura donc une bonne visualisation des clusters. Notons que ce type d'approche peut être implémenté dans une architecture parallèle, de par la définition des réseaux de neurones. Par contre, elles ne peuvent s'appliquer que sur des problèmes où les individus sont décrits par des attributs tous quantitatifs.

B.3.5 Les approches fondées sur le quadrillage de l'espace

Ces approches ont pour objectif de découper l'espace des données en segments (cubes, cellules, région) et ont été mises au point, principalement, pour la classification des données spatiales. La densité de chaque segment est calculée et les zones denses contiguës sont regroupées dans une même cluster. L'ensemble des clusters résultants est ensuite soumis à un algorithme de classification. Contrairement aux approches par partitionnement, telles que les méthodes \mathcal{K} -means, ces approches ne procèdent pas par réaffectation des individus. Elles sont plus proches des algorithmes hiérarchiques. Toutefois, l'agrégation des segments ne dépend pas d'une mesure de distance mais d'un paramètre fixé à l'avance.

La principale difficulté de ces méthodes est de choisir convenablement les paramètres liés à la taille des cellules. En effet, des cellules de trop grande taille reprendront beaucoup d'individus qui ne sont pas forcément bien uniformément répartis. On pourra donc avoir des clusters peu homogènes, nécessitant d'être subdivisés par la suite. Il y a donc sous-partitionnement. Par contre, des cellules de petite taille seront généralement très denses à partir du moment où elles contiennent des individus. On aura alors tendance à détecter des frontières qui n'auraient pas lieu d'être et à avoir un sur-partitionnement.

Les algorithmes CLIQUE (CLustering In QUEst)[[Agrawal et al., 1998](#)] et STING (STatistical INformation Grid-based method)[[Wang et al.](#)] sont apparentés à cette famille d'approches basées sur le quadrillage de l'espace des données et peuvent aussi être utilisés pour classer des données numériques hautement multidimensionnelles.

B.4 Évaluation d'une classification

B.4.1 Détermination du meilleur nombre de clusters

La recherche du meilleur nombre de clusters (ou classes) dans une partition est une phase très importante dans la classification, mais elle est souvent difficile et délicate. Dans les méthodes de classification par partition, le nombre de clusters à découvrir est un paramètre d'entrée à fixer par l'utilisateur. Ce qui n'est toujours pas évident. En effet, si le choix du nombre de clusters dans une partition ne correspond pas au meilleur nombre de clusters existantes dans un ensemble d'individus,

alors la partition choisie soit regroupe à tort des classes séparées (pas assez de clusters), soit découpe en plusieurs clusters des clusters homogènes (trop de clusters). Les méthodes de classification par partition proposent alors de fixer plusieurs valeurs de ce nombre de classes et les partitions correspondantes sont comparées.

La recherche du meilleur nombre de clusters se fait alors par essais et erreurs. Une même technique est utilisée à plusieurs reprises avec un nombre croissant de clusters, et pour chaque nouvelle partition obtenue, on calcule la valeur d'un critère de qualité. Le nombre de classes retenu est celui qui conduit à la meilleure valeur de ce critère. Dans cette section, nous passons en revue les différents critères permettant de trouver une valeur sinon optimale, du moins fiable de ce meilleur nombre de clusters. Nous décrivons alors dans les sous-sections de la section B.4.1 les indices les plus répandus pour la détermination du meilleur nombre de clusters. D'autre part, on peut également envisager d'utiliser des visualisations graphiques pour déterminer le meilleur nombre de clusters. En effet, lorsque la valeur du critère évolue de façon monotone avec le nombre de clusters et tend vers une asymptote horizontale lorsque le nombre de clusters augmente, la visualisation graphique permet de trouver dans la courbe représentant l'évolution de cette quantité un "coude" après lequel son évolution devient négligeable. Ce *coude* représente alors le meilleur nombre de clusters.

B.4.1.1 Indices fondés sur la somme de carrés

Soit \mathcal{K} le nombre de clusters formés à l'étape d'agrégation en cours. L'étape suivante fournira donc $\mathcal{K} - 1$ clusters.

a) R-Square (RSQ)

Ce critère (voir équation B.39) représente la proportion de variance expliquée par la partition en \mathcal{K} clusters. Ce qui correspond à l'inertie inter-cluster des \mathcal{K} clusters selon le critère de Ward (cf. équation B.38).

$$RSQ = R^2 = 1 - \frac{W}{T} \quad (\text{B.39})$$

$RSQ = 1$ lorsque toutes les observations forment des clusters individuels et diminuent progressivement jusqu'à la valeur 0 lorsque toutes les individus sont dans un seul cluster. Le *meilleur* nombre de clusters à retenir correspond à l'endroit sur la courbe où RSQ fait un saut important lorsque deux clusters sont agrégés.

b) Semi-Partial R-Square ($SPRSQ$)

Ce critère (voir équation B.40) représente la perte d'homogénéité quand deux clusters sont agrégés. Elle est mesurée par la réduction du RSQ avant et après

l'agrégation des clusters L_k et L_l , ce qui correspond à la perte d'inertie selon le critère de Ward.

$$SPRSQ = \frac{\mathbf{B}_{k \cup l}}{\mathbf{T}} \quad \text{où } \mathbf{B}_{k \cup l} = \mathbf{W}_{k \cup l} - \mathbf{W}_k - \mathbf{W}_l \quad (\text{B.40})$$

Dans ce critère, le nombre de clusters à retenir est donné par l'endroit où on note une forte réduction de $SPRSQ$. Donc, la stratégie consiste à choisir le meilleur nombre de clusters avec un *fort* RSQ et un *saut important* du $SPRSQ$ à l'agrégation suivante.

c) Root-Mean-Square Standard Deviation ($RMSSTD$)

Le $RMSSTD$ est la racine carrée de la variance du nouveau cluster $L_{k \cup l}$ formé par agrégation des clusters L_k et L_l , calculé comme la somme des variances de toutes les variables.

$$RMSSTD = \left[\frac{\mathbf{W}_{k \cup l}}{p(n_{k \cup l} - 1)} \right]^{1/2} \quad (\text{B.41})$$

Plus deux clusters agrégés se ressemblent, plus homogène est le cluster $L_{k \cup l} = L_k \cup L_l$ et donc, plus petit sera $RMSSTD$.

B.4.1.2 Indices fondés sur des pseudo-statistiques

a) La pseudo-statistique de [Calinski et Harabasz \[1974\]](#)

La pseudo-statistique de [Calinski et Harabasz \[1974\]](#), encore appelée **pseudo-F**, permet de comparer l'homogénéité d'une partition en k clusters à celle de la partition en $(k - 1)$ clusters. Elle est donnée par l'équation [B.42](#) suivante.

$$\text{pseudo-F}(k) = \frac{(\mathbf{T} - \mathbf{W}) / (k - 1)}{\mathbf{W} / (n - k)} \quad (\text{B.42})$$

La valeur de k qui maximise la pseudo-F(k) est choisie comme le meilleur nombre de clusters. Cette valeur se visualise graphiquement par un *pic* qui apparaît sur la courbe lorsque l'on représente sur un graphique les valeurs du pseudo-F en fonction du nombre de classes.

b) La pseudo-statistique T-carrée

La pseudo-statistique T-carrée est une statistique de type de [Hotelling \[1931\]](#) qui permet de comparer les vecteurs des moyennes de deux clusters. Si les moyennes

dans les deux clusters sont très différentes, alors les clusters ne doivent pas être agrégés. Elle est donnée par l'équation B.43 suivante.

$$\text{pseudo-T}^2(k) = \frac{\mathbf{B}_{kUl}}{(\mathbf{W}_k + \mathbf{W}_l)/(n_k + n_l - 2)} \quad \text{où } \mathbf{B}_{kUl} = \mathbf{W}_{kUl} - \mathbf{W}_k - \mathbf{W}_l \quad (\text{B.43})$$

Le meilleur nombre de clusters sera choisi de la manière suivante : faible pseudo-T² et fort pseudo-T² à l'agrégation suivante.

c) Cubic Clustering Criterion (CCC)

Ce critère, proposé par Publishing [1998], a été développé pour tester statistiquement l'hypothèse que les clusters sont des hyper-cubes obtenus à partir d'une distribution uniforme sur une hyper-boîte. Les grosses valeurs de CCC (CCC > 3) indiquent une bonne partition. Le Cubic Clustering Criterion teste les deux hypothèses suivantes :

H_0 : les données ont été échantillonnées d'une distribution uniforme sur une hyper-boîte.

H_1 : les données ont été échantillonnées d'une mixture de distributions normales sphériques multivariées, avec des variances et des probabilités d'échantillonnage égales.

Les valeurs positives de CCC signifient que le R^2 obtenu est plus élevé que celui attendu si l'échantillonnage était issue d'une distribution uniforme (delà, H_0 est rejeté). CCC est donc basé sur l'hypothèse qu'une distribution uniforme sur un hyper-cube de l'espace peut être divisée en clusters de forme grossière comme des hyper-cubes identiques. CCC est donné par l'équation B.44 suivante :

$$\text{CCC} = \log \left(\frac{1 - E(R^2)}{1 - R^2} \right) \frac{\sqrt{n\kappa/2}}{(0.001 + E(R^2))^2} \quad (\text{B.44})$$

où κ représente la valeur de la dimension inférieure ou égale à p (nombre de variables), $E(\cdot)$ représente l'espérance mathématique, R^2 est donné par l'équation B.39. Le critère CCC doit être utilisé dans le cadre des méthodes basées sur la minimisation des variances comme la méthode de Ward ou celle des k-means. Il est utilisé dans de bonnes conditions lorsque les clusters ont plus de 10 individus.

Un pic du CCC plus grand que 2 ou 3 indique une bonne classification tandis qu'un pic du CCC entre 0 et 2 indique que la classification doit être interprétée avec précaution. D'autre part, un pic précédé d'une forte montée du CCC, et suivi d'une descente plus douce, révèle une classification en clusters bien distinctes (soit 5 clusters). Les valeurs négatives et décroissantes pour des classifications de plus de deux clusters indiquent que, soit la distribution est unimodale, soit il y a une longue

queue de distribution. De fortes valeurs négatives (-30 par exemple) peuvent être dues à des points atypiques qu'il faudra alors enlever. Lorsque CCC continue à croître avec le nombre de clusters, il peut indiquer la formation de *poches* de sous-clusters dans un cluster *granuleux* (des points isolés par exemple).

B.4.1.3 L'indice de Goodman et Kruskal [1954]

Pour calculer cet indice, il faut examiner tous les quadruplets d'individus et les distances u induites par la partition. Si deux individus \mathbf{x}_i et \mathbf{x}_j appartiennent au même cluster de la partition, alors $u(\mathbf{x}_i, \mathbf{x}_j) = 0$ sinon $u(\mathbf{x}_i, \mathbf{x}_j) = 1$. Un quadruplet est dit cohérent si $u(\mathbf{x}_i, \mathbf{x}_j) < u(\mathbf{x}_k, \mathbf{x}_l)$ et $d(\mathbf{x}_i, \mathbf{x}_j) < d(\mathbf{x}_k, \mathbf{x}_l)$. Il est dit incohérent si $[u(\mathbf{x}_i, \mathbf{x}_j) < u(\mathbf{x}_k, \mathbf{x}_l)$ et $d(\mathbf{x}_i, \mathbf{x}_j) > d(\mathbf{x}_k, \mathbf{x}_l)]$ ou $[u(\mathbf{x}_i, \mathbf{x}_j) > u(\mathbf{x}_k, \mathbf{x}_l)$ et $d(\mathbf{x}_i, \mathbf{x}_j) < d(\mathbf{x}_k, \mathbf{x}_l)]$. Autrement dit, un quadruplet est dit cohérent (respectivement, incohérent) si la distance entre deux individus d'un même cluster est strictement plus petite (respectivement, grande) que la distance entre deux individus appartenant à des clusters différents. L'indice de Goodman et Kruskal [1954], est donné par l'équation B.45 suivante :

$$GK(k) = \frac{\mathcal{S}^+ - \mathcal{S}^-}{\mathcal{S}^+ + \mathcal{S}^-} \quad (\text{B.45})$$

où \mathcal{S}^+ représente le nombre de quadruplets cohérents et \mathcal{S}^- le nombre de quadruplets incohérents.

$GK(k) \in [-1; 1]$ et le meilleur nombre de clusters est donné par la valeur de k qui maximise $GK(k)$. Toutefois, on peut remarquer qu'un certain nombre de quadruplets sont inclassables à cause des relations d'égalité entre paires de distances.

B.4.1.4 La C -index ou indice de Hubert et Levin [1976]

La C -index est définie par l'équation B.46 suivante.

$$HL(k) = \frac{\mathcal{D}_{\text{intra}}(k) - r\mathcal{D}_{\text{min}}}{r\mathcal{D}_{\text{max}} - r\mathcal{D}_{\text{min}}} \quad (\text{B.46})$$

où $\mathcal{D}_{\text{intra}}(k)$ représente la somme des dissimilarités intra-clusters et si la partition a un total de r dissimilarités, \mathcal{D}_{min} (respectivement \mathcal{D}_{max}) est définie comme étant la somme des r plus petite (respectivement plus large) dissimilarités. $HL(k) \in [0; 1]$ et le meilleur nombre de clusters est donné par la valeur qui minimise $HL(k)$.

B.4.1.5 L'indice Silhouette de Rousseeuw [1987]

Pour chaque $\mathbf{x}_i \in \Omega$ tel que $\mathbf{x}_i \in L_k$ (avec $n_k = |L_k|$), on pose :

$$- a(\mathbf{x}_i) = \frac{1}{n_k - 1} \sum_{\mathbf{x}_j \in L_k - \{\mathbf{x}_i\}} \delta(\mathbf{x}_i, \mathbf{x}_j).$$

$$\begin{aligned}
- \forall l \in \{1, \dots, \mathcal{K}\} - \{k\}, \quad \mathcal{D}(\mathbf{x}_i, L_l) &= \frac{1}{n_l} \sum_{\mathbf{x}_j \in L_l} \delta(\mathbf{x}_i, \mathbf{x}_j). \\
- b(\mathbf{x}_i) &= \mathcal{D}(\mathbf{x}_i, L_m) = \min_{l \neq k} \mathcal{D}(\mathbf{x}_i, L_l).
\end{aligned}$$

La largeur de la silhouette $s(\mathbf{x}_i)$ est donnée par l'équation B.47 suivante.

$$s(\mathbf{x}_i) = \frac{b(\mathbf{x}_i) - a(\mathbf{x}_i)}{\max \{a(\mathbf{x}_i), b(\mathbf{x}_i)\}} \quad \text{avec } \mathbf{x}_i \in \Omega. \quad (\text{B.47})$$

avec $s(\mathbf{x}_i) \in [-1; +1]$ tels que :

- si $s(\mathbf{x}_i) \cong 1$, alors \mathbf{x}_i est bien classé dans L_k .
- si $s(\mathbf{x}_i) \cong 0$, alors \mathbf{x}_i est entre les clusters L_k et L_m .
- si $s(\mathbf{x}_i) \cong -1$, alors \mathbf{x}_i est mal classé (c'est-à-dire, \mathbf{x}_i plus proche de L_m que de L_k).

La qualité d'une partition \mathcal{P}_k de k clusters est la moyenne globale des largeurs des silhouettes dans les différents clusters L_r qui composent \mathcal{P}_k . L'indice silhouette $S(k)$ est donc donné par l'équation (B.48) suivante :

$$S(k) = \frac{1}{n} \sum_{r=1}^{\mathcal{K}} n_r \bar{S}_r = \frac{1}{n} \sum_{i=1}^n s(\mathbf{x}_i) \quad \text{où } \mathbf{x}_i \in \Omega. \quad (\text{B.48})$$

Où $\bar{S}_r = \frac{1}{n_r} \sum_{\mathbf{x}_i \in L_r} s(\mathbf{x}_i) \forall r \in \{1, \dots, \mathcal{K}\}$. $S(k) \in [-1; +1]$ et le meilleur nombre de clusters est donné par la valeur de k qui maximise $S(k)$.

B.4.1.6 L'indice de Davies et Bouldin [1979]

L'indice de Davies et Bouldin [1979] (voir équation (B.49)) est définie en fonction du rapport entre la somme des moyennes des distances intra-cluster et la distance entre les centres des clusters.

$$DB(k) = \frac{1}{k} \sum_{r,l=1}^k \max_{i \neq j} \frac{\mathcal{D}_{\text{moy}}(L_r) + \mathcal{D}_{\text{moy}}(L_l)}{d(g_r, g_l)} \quad (\text{B.49})$$

où $\mathcal{D}_{\text{moy}}(L_r) = \frac{1}{n_r} \sum_{\mathbf{x}_i, \mathbf{x}_j \in L_r} d(\mathbf{x}_i, \mathbf{x}_j)$ et représente la moyenne des distances entre les individus du cluster L_r .

Le meilleur nombre de clusters est donné par la valeur de k qui minimise l'indice de Davies et Bouldin [1979].

B.4.1.7 L'indice de Krzanowski et Lai's [1988]

Cet indice est défini par l'équation (B.50) suivante :

$$KL(k) = \left| \frac{(k-1)^{2/p} \mathbf{W}^{k-1} - (k)^{2/p} \mathbf{W}^k}{(k)^{2/p} \mathbf{W}^k - (k+1)^{2/p} \mathbf{W}^{k+1}} \right| \quad (\text{B.50})$$

où p représente le nombre de variables. Le meilleur nombre de clusters est donné par la valeur de k qui maximise $KL(k)$.

B.4.1.8 La connectivité

Soit $n_{i(j)}$ le j ème voisin le plus proche de l'individu \mathbf{x}_i , et $\mathbf{x}_{i,n_{i(j)}} = 0$ si \mathbf{x}_i et \mathbf{x}_j sont dans le même cluster et $1/j$ sinon. Alors, pour une partition particulière $\mathcal{P} = \{L_1, L_2, \dots, L_{\mathcal{K}}\}$ des n individus en \mathcal{K} clusters disjoints. La connectivité est définie par :

$$\text{Conn}(k) = \sum_{i=1}^n \sum_{j=1}^p \mathbf{x}_{i,n_{i(j)}}. \quad (\text{B.51})$$

$\text{Conn}(k) \in [0; \infty]$ et le meilleur nombre de clusters est donné par la valeur de k qui minimise $\text{Conn}(k)$.

B.4.1.9 L'indice de Dunn

L'indice de Dunn est le ratio de la plus petite distance entre les individus, pas du même cluster, à la plus large distance intra-cluster. Cet indice est calculé selon l'équation (B.52).

$$D(\mathcal{P}) = \frac{\min_{L_k, L_l \in \mathcal{P}, L_k \neq L_l} \left(\min_{\mathbf{x}_i \in L_k, \mathbf{x}_j \in L_l} \delta(\mathbf{x}_i, \mathbf{x}_j) \right)}{\max_{L_m \in \mathcal{P}} \text{diam}(L_m)} \quad (\text{B.52})$$

où $\text{diam}(L_m)$ est la distance maximum entre les individus du cluster L_m . $D(\mathcal{P}) \in [0; \infty]$, et le meilleur nombre de clusters est donné par la partition qui maximise $D(\mathcal{P})$.

B.4.1.10 La statistique Tibshirani *et al.* [2001]

Pour c dans l'intervalle $[1, \mathcal{K}]$

- Calculer la matrice de dispersion intra-cluster \mathbf{W}_c et son déterminant $\det(\mathbf{W}_c)$
- Construire la courbe des $\log[\det(\mathbf{W}_c)]$ en fonction du nombre c de clusters (avec $c = \{1, \dots, \mathcal{K}\}$)
- Comparer cette courbe à celle obtenue à partir des données uniformément réparties.

L'estimation du nombre de clusters à retenir est la valeur c correspondant au plus grand écart entre les deux courbes, ce qui représente la statistique Gap.

B.4.1.11 L'indice de Hartigan [1975]

L'indice de Hartigan $H(k)$ est une F-statistique partielle permettant de tester s'il est intéressant d'ajouter un cluster aux k déjà existants. Pour la détermination

du meilleur nombre de clusters, on choisit le premier $k \geq 1$ tel que $H(k) \leq 10$.

$$H(k) = (n - k - 1) \left(\frac{W^k}{W^{k+1}} - 1 \right) \quad (\text{B.53})$$

B.4.2 Validation des partitions

La validation d'une structure générée par une classification non supervisée est indispensable. Il existe plusieurs critères utilisés pour étudier la validation d'une classification, et les approches de validation qui se trouvent dans la littérature se regroupent en deux grands types : *les approches de validation dans un cadre probabiliste* et *les approches de validation dans un cadre non probabiliste*. Nous présentons dans cette thèse un résumé des approches de validation des clusters. Toutefois, le lecteur intéressé pourra consulter le livre de Jain and Dubes (1988), qui contient de nombreuses autres références.

B.4.2.1 Les critères de validation

Un critère de validation exprime la stratégie selon laquelle une structure de classification est validée. On distingue trois types de critères de validation, selon que l'on dispose ou pas d'information a priori sur les données : le critère interne, le critère externe et le critère relatif.

Le critère de validation externe

Le critère externe mesure le degré avec lequel les données confirment des informations connues *a priori* sur les données. Il permet aussi de comparer les résultats d'une classification non supervisée à une information sur la structure des données connue *a priori*.

Le critère de validation interne

Le critère interne mesure l'écart entre la structure engendrée par un algorithme de classification et les données, en tenant compte du biais introduit par l'utilisation d'un algorithme pour obtenir la structure de classification.

Le critère de validation relatif

Le critère relatif permet de comparer deux structures de partition. Il décide quelle structure est meilleure dans le sens plus stable ou mieux appropriée pour les données (voir équations B.46 et B.49).

B.4.2.2 La validation dans un cadre non probabiliste

Les approches de validation dans un cadre non probabiliste peuvent se répartir en trois grandes approches.

a) La première approche : mesure de l'adéquation

Elle permet de mesurer l'adéquation des résultats avec la dissimilarité initiale. L'idée consiste alors à construire un indice mesurant le lien entre la structure et les données initiales, puis valider la valeur de cet indice. Pour valider cet indice, on teste l'hypothèse nulle H_0 d'absence de structure en clusters [Bailey et Dubes, 1982] grâce à un test appelé *test de Monte Carlo* [Jain et Dubes, 1988]. Ensuite, on simule des données selon H_0 et puis on calcule la valeur de l'indice qui évalue la structure de classification générée par la méthode de classification utilisée sur les données initiales. Enfin, on teste si l'indice obtenu sur les données initiales est en adéquation avec les valeurs obtenues sur les données simulées.

b) La deuxième approche : mesure de la stabilité

Elle permet de mesurer la stabilité des résultats obtenus. L'idée consiste à construire (et valider) un indice mesurant la stabilité des résultats de la classification obtenue par rapport aux différentes perturbations que les données peuvent subir. Cet indice est défini par l'écart entre la structure initiale et la structure obtenue sur les données bruitées ou par la variation d'un critère mesuré sur ces deux structures [Rand; Gordon, 1999; Milligan, 1996]. Mesurer la stabilité d'une classification revient donc à étudier le comportement de la structure initiale face aux différentes perturbations tels que le fait de retrait d'un élément de l'ensemble des données ou la perturbation de la matrice de proximité.

c) La troisième approche : la validation croisée

Elle mesure l'écart entre les classifications obtenues sur un échantillon d'apprentissage et sur un échantillon test. L'idée de cette troisième approche consiste à construire (et à valider) un indice mesurant l'écart entre les classifications obtenues sur un échantillon d'apprentissage et sur un échantillon test. Le principe est proche de la validation croisée : on divise l'échantillon de base en deux parties Ω_1 et Ω_2 ; on applique ensuite une méthode de classification à chacun des deux échantillons ; puis, on mesure l'écart entre la partition \mathcal{P}_2 de Ω_2 générée par la méthode de classification, à celle obtenue en affectant les éléments de Ω_2 à la partition \mathcal{P}_1 de Ω_1 , en utilisant une règle d'affectation. Plus cet écart est faible, plus la partition générée sur l'ensemble tout entier est valide [voir Mcintyre et Blashfield, 1980; Smith et Jain, 1984; Breckenridge, 1989].

B.4.2.3 La validation dans un cadre probabiliste

La validation dans un cadre probabiliste pose généralement trois grands problèmes : la classifiabilité des données, le nombre de classes, et la stabilité des résultats, où il s'agit de déterminer si les résultats sont de même nature sur d'autres échantillons issus de la même famille de loi de probabilité. La plupart des tests statistiques sur le bien-fondé d'une partition s'appuient sur la loi limite (lorsque le nombre de l'échantillon tend vers l'infini) de statistiques sous certaines hypothèses de classifiabilité et de non classifiabilité. On distingue deux approches différentes, essentielles pour ces problèmes de validation : l'une par des outils descriptifs, graphiques et empiriques ; l'autre par des tests d'hypothèse dans les statistiques inductives. Bock [cf. [Arabie et al., 1996](#)] a proposé des propriétés de ces tests significatifs pour établir la différence entre l'hypothèse d'homogénéité d'une population et l'hypothèse alternative de classification ou d'hétérogénéité.

B.4.3 Comparaison des partitions

L'objectif de la comparaison des partitions est plus précisément d'évaluer une structure générée par une classification non supervisée en k clusters (partition \mathcal{P}_k), obtenue à partir d'un ensemble Ω (en utilisant un algorithme de classification autre que hiérarchique), en la comparant à une partition indépendante \mathcal{P}_l en l clusters obtenues sur un même ensemble d'individus Ω . La comparaison des partitions est basée sur la comparaison des clusters qui forment les partitions. La comparaison des clusters, quand à elle, est effectuée sur la base de leur *extension*, c'est-à-dire, sur l'ensemble d'individus affectés aux clusters. Cette stratégie s'oppose aux techniques de comparaison de partition basées sur l'*intension*, c'est-à-dire, sur l'ensemble de représentants (réels ou fictifs) des clusters [[Gomes Da Silva, 2009](#); [Nakache et Confais, 2004](#)].

Soient \mathcal{P}_k et \mathcal{P}_l les partitions résultantes de deux classifications sur un même ensemble d'individus Ω . La matrice de contingence sur ces deux partitions est donnée par la Table B.3. La valeur n_{ij} dans chaque cellule de ce tableau indique le nombre d'individus présents simultanément dans le cluster L_i de la partition \mathcal{P}_k et dans le cluster L_j de la partition \mathcal{P}_l . La partition \mathcal{P}_k est considérée la partition de référence.

B.4.3.1 La F-mesure

Le tableau de contingence (Table B.3) nous permet de définir les concepts de Rappel et Précision selon la formule donnée par l'équation B.54. Le rappel est défini par le ratio entre le nombre d'individus en commun entre deux clusters de différentes partitions et le total d'individus dans le cluster de la partition \mathcal{P}_k . La précision est définie par le ratio entre le nombre d'individus en commun entre deux

Clusters de $\mathcal{P}_k \backslash$ Clusters de \mathcal{P}_l	L_1	L_2	\dots	L_j	\dots	L_l	Total
L_1	n_{11}	n_{12}	\dots	n_{1j}	\dots	n_{1l}	$n_{1\bullet}$
L_2	n_{21}	n_{22}	\dots	n_{2j}	\dots	n_{2l}	$n_{2\bullet}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
L_i	n_{i1}	n_{i2}	\dots	n_{ij}	\dots	n_{il}	$n_{i\bullet}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
L_k	n_{k1}	n_{k2}	\dots	n_{kj}	\dots	n_{kl}	$n_{k\bullet}$
Total	$n_{\bullet 1}$	$n_{\bullet 2}$	\dots	$n_{\bullet j}$	\dots	$n_{\bullet l}$	$n_{\bullet\bullet} = n$

TABLE B.3 – Tableau de contingence entre les deux partitions \mathcal{P}_k et \mathcal{P}_l

clusters de différentes partitions et le total d'individus dans le cluster de la partition \mathcal{P}_l .

$$\text{Rappel}(L_i, L_j) = \frac{n_{ij}}{n_{i\bullet}} \quad \text{et} \quad \text{Précision}(L_i, L_j) = \frac{n_{ij}}{n_{\bullet j}} \quad (\text{B.54})$$

La F-mesure [Rijsbergen, 1979] combine les concepts de Rappel et Précision en un seule formule donnée par l'équation B.55. La formulation générale (cf. équation B.56) est donnée par la F_β -mesure. Pour $\beta = 1$, la F-mesure correspond à la moyenne harmonique entre Rappel et Précision.

$$F\text{-mesure}(L_i, L_j) = \frac{2 \cdot \text{Rappel}(L_i, L_j) \cdot \text{Précision}(L_i, L_j)}{\text{Rappel}(L_i, L_j) + \text{Précision}(L_i, L_j)} \quad (\text{B.55})$$

$$F_\beta\text{-mesure}(L_i, L_j) = \frac{(1 + \beta^2) \cdot \text{Rappel}(L_i, L_j) \cdot \text{Précision}(L_i, L_j)}{\text{Rappel}(L_i, L_j) + \beta^2 \cdot \text{Précision}(L_i, L_j)} \quad (\text{B.56})$$

La F-mesure vise à trouver le meilleur représentant d'un cluster L_i (avec $i \in \{1, \dots, k\}$) dans la partition \mathcal{P}_k par un cluster L_j (avec $j \in \{1, \dots, l\}$) dans la partition \mathcal{P}_l . Si pour (i, j) quelconque on a $\text{Rappel}(L_i, L_j) = \text{Précision}(L_i, L_j) = 1$, alors $L_i = L_j$. Les deux clusters L_i et L_j sont d'autant plus semblables lorsque leurs valeurs du Rappel et de la Précision sont très élevées. Si la valeur du Rappel est petite (respectivement, grande), alors cela signifie que le cluster L_i de la partition \mathcal{P}_k a perdu beaucoup (respectivement, peu) d'individus vis-à-vis de son cluster correspondant dans la partition \mathcal{P}_l . Si la valeur de la Précision est petite (respectivement, grande), alors cela signifie que le cluster L_i a gagné beaucoup (respectivement, peu) d'individus au regard de son cluster correspondant dans la partition \mathcal{P}_k .

Au niveau de la partition, la F-mesure F est calculée à partir des moyennes des valeurs maximales sur les colonnes du tableau de contingence (cf. équation B.57).

$$F(\mathcal{P}_k, \mathcal{P}_l) = \frac{1}{n} \sum_{i=1}^k n_{i\bullet} \max_{j=1, \dots, l} F\text{-mesure}(L_i, L_j) \quad (\text{B.57})$$

La F -mesure F assume des valeurs dans le rang $[0; 1]$ et a pour but fournir une évaluation numérique de la similarité entre deux partitions issues d'un même ensemble d'individus Ω . Les deux partitions \mathcal{P}_k et \mathcal{P}_l sont d'autant plus similaires lorsque la valeur de leur F -mesure est très élevée.

B.4.3.2 L'indice de Rand

Le tableau de contingence B.3 nous permet de déduire les formules suivantes :

$$\binom{n_{ij}}{2} = \frac{n_{ij}(n_{ij} - 1)}{2} \text{ représente le nombre de paires d'individus de } L_i \cap L_j.$$

$$\binom{n}{2} = \frac{n(n - 1)}{2} \text{ représente le nombre de paires d'individus de } \Omega.$$

$$\binom{n_{i\bullet}}{2} = \frac{n_{i\bullet}(n_{i\bullet} - 1)}{2} \text{ représente le nombre de paires d'individus de } L_i.$$

$$\binom{n_{\bullet j}}{2} = \frac{n_{\bullet j}(n_{\bullet j} - 1)}{2} \text{ représente le nombre de paires d'individus de } L_j.$$

$$a = \sum_{i=1}^k \sum_{j=1}^l \binom{n_{ij}}{2} \text{ représente le nombre de paires d'individus de } \Omega \text{ qui se retrouvent dans le même cluster et dans chacune des deux partitions } \mathcal{P}_k \text{ et } \mathcal{P}_l.$$

$$b = \sum_{i=1}^k \binom{n_{i\bullet}}{2} - \sum_{i=1}^k \sum_{j=1}^l \binom{n_{ij}}{2} \text{ représente le nombre de paires d'individus de } \Omega \text{ qui se retrouvent dans un même cluster } L_i \text{ de la partition } \mathcal{P}_k \text{ et dans deux clusters différents de la partition } \mathcal{P}_l.$$

$$c = \sum_{j=1}^l \binom{n_{\bullet j}}{2} - \sum_{i=1}^k \sum_{j=1}^l \binom{n_{ij}}{2} \text{ représente le nombre de paires d'individus de } \Omega \text{ qui se retrouvent dans un même cluster } L_j \text{ de la partition } \mathcal{P}_l \text{ et dans deux clusters différents de la partition } \mathcal{P}_k.$$

$$d = \binom{n}{2} + \sum_{i=1}^k \sum_{j=1}^l \binom{n_{ij}}{2} - \sum_{i=1}^k \binom{n_{i\bullet}}{2} - \sum_{i=1}^k \binom{n_{\bullet j}}{2} \text{ représente le nombre de paires d'individus de } \Omega \text{ qui se retrouvent dans deux clusters différents et dans les deux partitions } \mathcal{P}_k \text{ et } \mathcal{P}_l.$$

Les nombres a , b , c et d sont utilisés dans la définition de l'indice de Rand comme le montre l'équation (B.58) suivante.

$$\begin{aligned} R(\mathcal{P}_k, \mathcal{P}_l) &= \frac{a + d}{a + b + c + d} \\ &= \frac{\binom{n}{2} + \sum_{i=1}^k \sum_{j=1}^l \binom{n_{ij}}{2} - \left[\sum_{i=1}^k \binom{n_{i\bullet}}{2} + \sum_{j=1}^l \binom{n_{\bullet j}}{2} \right]}{\binom{n}{2}} \end{aligned} \quad (\text{B.58})$$

L'indice Rand, proposé par [Hubert et Arabie](#), mesure le pourcentage de concordance entre les partitions \mathcal{P}_k et \mathcal{P}_l . Il est symétrique, c'est-à-dire, $R(\mathcal{P}_k, \mathcal{P}_l) = R(\mathcal{P}_l, \mathcal{P}_k)$ et prend des valeurs dans l'intervalle $[0; 1]$. Si les deux partitions comparées sont identiques (c'est-à-dire que les deux partitions ont le même nombre de clusters et pour tout $r = 1, \dots, k$, il existe $s = 1, \dots, l$ tel que $L_r = L_s$) alors $R(\mathcal{P}_k, \mathcal{P}_l) = 1$.

B.4.3.3 L'indice de Rand Ajusté

L'indice de Rand Ajusté, proposé par [Hubert et Arabie](#), est donné par l'équation (B.59) suivante.

$$RA(\mathcal{P}_k, \mathcal{P}_l) = \frac{\sum_{i=1}^k \sum_{j=1}^l \binom{n_{ij}}{2} - \binom{n}{2}^{-1} \sum_{i=1}^k \binom{n_{i\bullet}}{2} \sum_{j=1}^l \binom{n_{\bullet j}}{2}}{\frac{1}{2} \left[\sum_{i=1}^k \binom{n_{i\bullet}}{2} + \sum_{j=1}^l \binom{n_{\bullet j}}{2} \right] - \binom{n}{2}^{-1} \sum_{i=1}^k \binom{n_{i\bullet}}{2} \sum_{j=1}^l \binom{n_{\bullet j}}{2}} \quad (\text{B.59})$$

L'indice RA assume des valeurs contenues dans l'intervalle $[-1; +1]$. Ainsi comme la F -measure, les valeurs proches de 1 correspondent à des partitions très semblables. Les valeurs proches de 0 sont vérifiées lorsque les accords entre les deux partitions sont dues au hasard, et les valeurs négatives lorsque les partitions sont peu liées.

B.4.4 Caractérisation des clusters

Nous présentons dans cette sous-section quelques critères qui peuvent être adoptés pour l'interprétation des clusters découverts par la méthode de classification.

B.4.4.1 Caractérisation unidimensionnelle

Pour déterminer les variables les plus caractéristiques de chaque cluster, on compare la moyenne (ou la fréquence) d'une variable dans le cluster à la moyenne (ou la fréquence) de cette variable dans l'échantillon total, en faisant l'hypothèse nulle H_0 que les individus qui constituent le cluster sont tirés au hasard et sans remise de l'échantillon global.

a) Variable nominale

On s'intéresse à la variable aléatoire N représentant le nombre d'individus du cluster L_k présentant la modalité e_j de la variable nominale X_j . Sous l'hypothèse H_0 , cette variable suit une distribution hypergéométrique de moyenne $E_k(N) = n_k \frac{n_{e_j}}{n}$ et de variance $\sigma_k^2(N) = n_k \cdot \frac{n-n_k}{n-1} \cdot \frac{n_{e_j}}{n} \left(1 - \frac{n_{e_j}}{n}\right)$ où $n = |\Omega|$, $n_k = |L_k|$ et

n_{e_j} est le nombre d'individus de Ω présentant la modalité e_j de la variable nominale X_j .

En pratique, cette distribution peut être approximée par une distribution normale si les effectifs des clusters sont assez élevés. Dans ce cas,

$$t_k(N) = \frac{N - E_k(N)}{\sigma_k(N)} \rightsquigarrow \mathcal{N}(0; 1) \quad (\text{B.60})$$

Elle représente l'écart entre la proportion dans le cluster L_k et la proportion globale en nombre d'écart-type d'une loi normale. Ainsi, le degré de significativité p du test de l'hypothèse nulle H_0 : "les nombres d'individus présentant la modalité e_j de la variable nominale X_j dans le clusters L_k et dans l'échantillon global sont égaux (au fluctuation aléatoire près)" est donné par :

$$p_k(e_j) = P[|X| > t_k(N_{ke_j})] \quad (\text{B.61})$$

où X est une variable normale centrée réduite et N_{ke_j} est le nombre observé d'individus présentant la modalité e_j dans le cluster L_k . Plus cette probabilité $p_k(e_j)$ est faible, plus l'hypothèse d'un tirage au hasard est rejetée et plus la modalité e_j est caractéristique du cluster L_k .

b) Variable continue

Sous la même hypothèse nulle H_0 : "les individus du cluster L_k sont tirés au hasard et sans remise de l'échantillon global Ω ". Ainsi, sous H_0 , les moyennes de la variable continue X_l dans le cluster L_k (\bar{X}_l^k) et dans l'échantillon global (\bar{X}_l) sont égales aux fluctuations aléatoires près. La valeur test est donnée par l'équation B.62 suivante :

$$t_k(X_l) = \frac{\bar{X}_l^k - \bar{X}_l}{\sigma_k(\bar{X}_l)} \quad \text{avec} \quad \sigma_k^2(\bar{X}_l) = \frac{n - n_k}{n - 1} \cdot \frac{\sigma^2(\bar{X}_l)}{n_k} \quad (\text{B.62})$$

où $\sigma^2(\bar{X}_l)$ est la variance de la variable continue X_l calculée dans l'échantillon global. Plus la valeur test $t_k(X_l)$ est grande, plus la moyenne de X_l dans le cluster L_k diffère de la moyenne globale, et plus la variable est caractéristique du cluster L_k .

B.4.4.2 Caractérisation multidimensionnelle

La notion de décomposition d'inertie, proposée par [Celeux et al. \[1989\]](#), peut être utilisée pour l'interprétation des clusters. Le but étant de repérer les variables les plus discriminantes dans la construction de chaque cluster.

a) Décomposition de l'inertie sur les clusters et les variables

Les inerties peuvent être décomposées en fonction des clusters qui constituent la partition ou encore en fonction des variables qui décrivent les individus.

L'inertie totale \mathbf{T} sur l'ensemble d'individus Ω peut s'écrire comme suit :

$$\mathbf{T} = \sum_{k=1}^{\mathcal{K}} \sum_{j=1}^p \mathbf{T}_k^j \quad \text{avec } \mathbf{T}_k^j = \sum_{\mathbf{x}_i \in L_k} \lambda_i d^2(x_i^j, g^j) \quad (\text{B.63})$$

\mathbf{T}_k^j représente l'écart moyen, sur la variable X_j , des individus du cluster L_k au centre de gravité du nuage de points.

L'inertie intra-cluster \mathbf{W} associée à la partition $\mathcal{P} = \{L_1, L_2, \dots, L_{\mathcal{K}}\}$ peut s'écrire comme dans l'équation (B.64) suivante :

$$\mathbf{W} = \sum_{k=1}^{\mathcal{K}} \sum_{j=1}^p \mathbf{W}_k^j \quad \text{avec } \mathbf{W}_k^j = \sum_{\mathbf{x}_i \in L_k} \lambda_i d^2(x_i^j, g_k^j) \quad (\text{B.64})$$

\mathbf{W}_k^j correspond à l'inertie intra-cluster du cluster L_k au niveau de la variable X_j .

L'inertie inter-cluster \mathbf{B} associée à la partition $\mathcal{P} = \{L_1, L_2, \dots, L_{\mathcal{K}}\}$ peut s'écrire comme dans l'équation (B.65) suivante :

$$\mathbf{B} = \sum_{k=1}^{\mathcal{K}} \sum_{j=1}^p \mathbf{B}_k^j \quad \text{avec } \mathbf{B}_k^j = \mu_k d^2(g_k^j, g^j) \quad (\text{B.65})$$

\mathbf{B}_k^j représente l'écart, sur la variable X_j , du centre de gravité du cluster L_k au centre de gravité du nuage de points.

De ce fait, nous pouvons décomposer les inerties sur les clusters et sur les variables séparément, comme le montre les équations suivantes.

– Décomposition des inerties sur les clusters

$$\mathbf{T}_k = \sum_{\mathbf{x}_i \in L_k} \lambda_i d^2(x_i^j, g) \quad (\text{B.66})$$

$$\mathbf{W}_k = \sum_{\mathbf{x}_i \in L_k} \lambda_i d^2(x_i^j, g_k) \quad (\text{B.67})$$

$$\mathbf{B}_k = \mu_k d^2(g_k^j, g) \quad (\text{B.68})$$

– Décomposition des inerties sur les variables

$$\mathbf{T}^j = \sum_{i=1}^n \lambda_i d^2(x_i^j, g^j) \quad (\text{B.69})$$

$$\mathbf{W}^j = \sum_{k=1}^{\mathcal{K}} \sum_{\mathbf{x}_i \in L_k} \lambda_i d^2(x_i^j, g^j) \quad (\text{B.70})$$

$$\mathbf{B}^j = \sum_{k=1}^{\mathcal{K}} \mu_k d^2(g_k^j, g^j) \quad (\text{B.71})$$

Ces décompositions de l'inertie peuvent être liées de façon additive par la relation fondamentale $\mathbf{T} = \mathbf{W} + \mathbf{B}$ comme le montre les équations suivantes.

$$\sum_{k=1}^{\mathcal{K}} \sum_{j=1}^p \mathbf{T}_k^j = \sum_{k=1}^{\mathcal{K}} \sum_{j=1}^p \mathbf{W}_k^j + \sum_{k=1}^{\mathcal{K}} \sum_{j=1}^p \mathbf{B}_k^j \quad (\text{B.72})$$

$$\sum_{k=1}^{\mathcal{K}} \mathbf{T}_k = \sum_{k=1}^{\mathcal{K}} \mathbf{W}_k + \sum_{k=1}^{\mathcal{K}} \mathbf{B}_k \quad (\text{B.73})$$

$$\sum_{j=1}^p \mathbf{T}^j = \sum_{j=1}^p \mathbf{W}^j + \sum_{j=1}^p \mathbf{B}^j \quad (\text{B.74})$$

b) Contribution des variables

La contribution de l'inertie expliquée en groupant les individus de Ω en une partition $\mathcal{P} = (L_1, L_2, \dots, L_{\mathcal{K}})$ est donné par l'équation (B.75).

$$\mathbf{R} = \frac{\mathbf{B}}{\mathbf{T}} \quad (\text{B.75})$$

Plus la valeur de \mathbf{R} est élevée, plus la partition constitue une bonne représentation des individus. Lorsque $\mathbf{R} = 1$, le nuage d'individus correspond aux \mathcal{K} prototypes des clusters.

Pour chaque variable X_j , un indice analogue à \mathbf{R} est défini par l'équation B.76.

$$\mathbf{C}(\mathcal{X}_{\bullet j}) = \frac{\mathbf{B}^j}{\mathbf{T}^j} \quad \forall j = 1, \dots, p \quad (\text{B.76})$$

Cet indice représente la contribution d'inertie de la variable X_j pris en compte par la partition \mathcal{P} . Il mesure ainsi le pouvoir discriminant de la variable X_j par rapport à la partition \mathcal{P} . Une relation existe entre cet indice et \mathbf{R} (cf. équation (B.75)) :

- Si $\mathbf{C}(X_j) > \mathbf{R}$, la variable X_j est plus discriminante que la moyenne car \mathbf{R} représente le pouvoir discriminant moyen vis-à-vis de la partition \mathcal{P} .

– Si $C(X_j) < \mathbf{R}$, la variable X_j est moins discriminante que la moyenne. La *contribution relative* de la variable X_j à l'inertie inter-cluster de la partition \mathcal{P} est donnée par l'équation (B.77).

$$C_R(X_j) = \frac{\mathbf{B}^j}{\mathbf{B}} \quad \text{avec} \quad \sum_{j=1}^p C_R(X_j) = 1 \quad (\text{B.77})$$

B.4.4.3 Interprétation statistique des clusters par variable

Les indices d'interprétation des clusters par un ensemble de variables [Celeux *et al.*, 1989; Gomes Da Silva, 2009] sont fondés sur la décomposition de l'inertie associée à un ensemble d'individus et à la partition dans laquelle ceux-ci sont classés. La technique d'interprétation des clusters par des variables consiste à calculer, pour chacune des variables, un critère mesurant sa pertinence dans l'interprétation du cluster. Pour chaque variable X_j et chaque cluster L_k , on définit les critères suivants :

a) Pouvoir discriminant de la variable X_j par rapport au cluster L_k

Elle est donnée par la formule de l'équation (B.78) suivante :

$$C(X_j, L_k) = \frac{\mathbf{B}_k^j}{\mathbf{T}^j} \quad \text{avec} \quad C(X_j) = \sum_{k=1}^{\mathcal{K}} C(X_j, L_k) \quad (\text{B.78})$$

Cet indice représente le pourcentage du pouvoir discriminant de la variable X_j pris en compte par le cluster L_k . Plus cet indice est fort, plus la variable X_j aura un comportement homogène sur les individus du cluster L_k . L'examen de cet indice est d'un atout majeur pour les variables à fort pouvoir discriminant, il indique la répartition entre les clusters de ce pouvoir discriminant.

b) Contribution relative de la variable X_j à l'inertie intercluster du cluster L_k

Elle est donnée par la formule de l'équation (B.79) suivante :

$$C_R(X_j, L_k) = \frac{\mathbf{B}_k^j}{\mathbf{B}_k} \quad (\text{B.79})$$

Cet indice est complémentaire au précédent et permet de déterminer les variables qui caractérisent le plus chaque cluster.

Bibliographie

- B. ABEL-SMITH et P. TOWNSEND : *The Poor and the Poorest*. Bell and Sons, London, 1965.
- P. AGHION et P. BOLTON : A theory of trickle-down growth and development. *Review of Economic Studies*, 64(2):151–72, 1997.
- R. AGRAWAL, J. GEHRKE, D. GUNOPULOS et P. RAGHAVAN : Automatic subspace clustering of high dimensional data for data mining applications. *SIGMOD Rec.*, 27(2):94–105, 1998.
- S. ALKIRE : *Valuing Freedoms : Sen's Capability Approach and Poverty Reduction*. Oxford University Press, New York, 2005.
- S. ALKIRE et J. FOSTER : Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7-8):476–487, 2011.
- S. ALKIRE et M.E. SANTOS : Acute multidimensional poverty : A new index for developing countries. *Oxford Poverty & Human Development Initiative (OPHI)*, WORKING PAPER No. 38, 2010.
- R.M. ANDERBERG : *Cluster analysis for applications*. Academic Press, New York, 1973.
- M. ANKERST, M.M. BREUNIG, H.-P. KRIEGEL et J. SANDER : OPTICS : Ordering points to identify the clustering structure. pages 49–60. ACM Press, 1999.
- K. APPIAH-KUBI, E. AMANNING-AMPOMAH et C. AHORTOR : Multidimensional analysis of poverty in Ghana using fuzzy sets theory. Cahiers de recherche PMMA 2007-21, PEP-PMMA, 2007.
- P. ARABIE, L.J. HUBERT et G. De SOETE : *Clustering and classification*. World Scientific Publishing. River Edge, 1996.
- A. ARNAND et A.K. SEN : Multidimensional Analysis of Poverty in Ghana using Fuzzy Sets Theory. Human Development Papers, UNDP, New York, 1997.
- K.J. ARROW : *Social Choice and Individual Values*. 2nd Edition in 1963, John Wiley & Sons, New York, 1951.
- L.M. ASSELIN : *Analysis of Multidimensional Poverty : Theory and Case Studies*. Economic studies in inequality, social exclusion and well-being ; vol. 7. Springer, IDRC, New York, 2009.

- A.B. ATKINSON : On the measurement of poverty. *Econometrica*, 55(4):749–64, 1987.
- A.B. ATKINSON : Multidimensional deprivation : Contrasting social welfare and counting approaches. *Journal of Economic Inequality*, 1(1):51–65, 2003.
- A.B. ATKINSON et F. BOURGUIGNON : The comparison of multi-dimensioned distributions of economic status. *Review of Economic Studies*, 49(2):183–201, 1982.
- H. BACELAR-NICOLAU : The affinity coefficient in cluster analysis. *Methods on Operations Research*, 53:507–512, 1985.
- H. BACELAR-NICOLAU : Classifying integer scale data by the affinity coefficient. *Methods on Operations Research*, 60:587–595, 1990.
- Jr.T.A. BAILEY et R.C. DUBES : Cluster validity profiles. *Pattern Recognition*, 15(2):61–83, 1982.
- R. BAUER : *Social Indicators*. M.I.T. Press, Cambridge, MA, 1967.
- E. BÉNICOURT : *Les analyses du PNUD et de la Banque mondiale sur la pauvreté et le développement : la place d'Amartya Sen*. Thèse de doctorat, EHESS, Paris, 2005.
- J. BENTHAM : *An Introduction to the Principles of Morals and Legislation*. Oxford : Clarendon Press, London, 1789.
- J. BENTHAM : Article on utilitarianism. In A. Goldworth (Ed.), *Deontology together with A Table of the Springs of Action and the Article on Utilitarianism*. Oxford : Clarendon Press, London, pages 283–328, 1829.
- J. BENTHAM : *Parliamentary Candidate's Proposed Declaration of Principles*. Published at the office of the Westminster Review by Robert Heward, London, 1831.
- J.-P. BENZÉCRI : *Correspondence Analysis Handbook*. Marcel Dekker, New York, 1992.
- A. BERGSON : A reformulation of certain aspects of welfare economics. *The Quarterly Journal of Economics*, 52(2):310–334, 1938.
- P. BERKHIN : Survey of clustering data mining techniques. Technical report, Accrue Software, 2002.

- A. BERTIN : *Pauvreté Monétaire, Pauvreté non Monétaire : une Analyse des Interactions Appliquée à la Guinée*. Thèse de doctorat, Université Montesquieu-Bordeaux IV, 2007.
- E. BÉNICOURT : Contre Amartya Sen. *L'économie Politique*, 3(23):72–82, 2004.
- E. BÉNICOURT : Amartya Sen : une nouvelle ère pour le développement ? *Revue Tiers Monde*, 47(186):433–447, 2006.
- C. BOOTH : *Life and Labour of the People of London*. Macmillan, London, New York, 1892.
- C. BOOTH : *The Aged Poor in England and Wales*. Macmillan, London, 1894.
- F. BOOYSEN, R. BURGER, G. Du RAND, M. Von MALTITZ et S. Van der BERG : Trends in poverty and inequality in seven African countries. Cahiers de recherche PMMA 2007-06, PEP-PMMA, 2007.
- I. BORG et P.J.F GROENEN : *Modern Multidimensional Scaling : Theory and Applications*. Springer, 2005.
- F. BOURGUIGNON et S.R. CHAKRAVARTY : Multi-dimensional poverty orderings. DELTA Working Papers 2002-22, DELTA (École Normale Supérieure), 2002.
- F. BOURGUIGNON et S.R. CHAKRAVARTY : The measurement of multidimensional poverty. *Journal of Economic Inequality*, 1(1):25–49, 2003.
- D. BOUYSSOU : Outranking relations : Do they have special properties ? *Journal of Multi-Criteria Decision Analysis*, 5(2):99–111, 1996.
- D. BOUYSSOU, T. MARCHANT, M. PIRLOT, P. PERNY, A. TSOUKIÀS et Ph. VINCKE : *Evaluation and Decision Models : a Critical Perspective*. Kluwer Academic, Boston, 2000.
- A. BRANDOLINI et G. D'ALESSIO : Measuring well-being in the functioning space. Working paper, Bank of Italia, Research Department, Mimeo, 1998.
- J.N. BRECKENRIDGE : Replicating cluster analysis : Method, consistency, and validity. *Multivariate Behavioral Research*, 24(2):147–161, 1989.
- T. CALINSKI et J. HARABASZ : A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods*, 3(1):1–27, 1974.
- G. CELEUX, E. DIDAY, G. GOVAERT, Y. LECHEVALLIER et H. RALAMBON-DRAINY : *Classification automatique des données : environnement statistique et informatique*. Dunod, Paris, 1989.

- A. CERIOLI et S. ZANI : A fuzzy approach to the measurement of poverty. *In C. Dagum and M. Zenga (eds.) Income and Wealth Distribution, Inequality and Poverty, Studies in Contemporary Economics, Springer Verlag, Berlin*, pages 272–284, 1990.
- S.R. CHAKRAVARTY, D. MUKHERJEE et R.R. RANADE : On the family of subgroup and factor decomposable measures of multidimensional poverty. *In Slottje D.J. (eds.), Research on Economic Inequality, Stanford, CT and London, JAI Press*, 8:175–194, 1998.
- M. CHAVENT : A monothetic clustering method. *Pattern Recognition Letters*, 19 (11):989–996, 1998.
- M. CHAVENT, Y. LECHEVALLIER et O. BRIANT : DIVCLUS-T : A monothetic divisive hierarchical clustering method. *Computational Statistics & Data Analysis*, 52(2):687–701, 2007.
- B. CHELI et A. LEMMI : A totally fuzzy and relative approach to the multidimensional analysis of poverty. *Economic Notes Monte dei Paschi di Siena*, 24 (1):115–134, 1995.
- S. CHEN et M. RAVALLION : The changing profile of poverty in the world. 2020 vision briefs, International Food Policy Research Institute (IFPRI), 2007.
- M.E. CHIAPPERO : A new approach to evaluation of well-being and poverty by fuzzy set theory. *Giornale degli Economisti e Annali di Economia*, 7-9(53):367–388, 1994.
- T. COELLI, D.S. PRASADA RAO et G.E. BATTESE : *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic, Boston, 1998.
- COMMONWEALTH : *Engendering Adjustment for the 1990s*. Report of a Commonwealth Group on Women and Structural Adjustment, Commonwealth Secretariat, London, 1989.
- G. CORNIA, R. JOLLY et F. STEWART : *L'ajustement à visage humain*. Economica-Unicef, Paris, 1987.
- F.A. COWELL : *Measuring Inequality*. Philip Allan (first edition), Oxford, 1977.
- F.A. COWELL et K. KUGA : Additivity and the entropy concept : An axiomatic approach to inequality measurement. *Journal of Economic Theory*, 25(1):131–143, 1981a.
- F.A. COWELL et K. KUGA : Inequality measurement : An axiomatic approach. *European Economic Review*, 15(3):287–305, 1981b.

- T.F. COX et M.A.A. COX : *Multidimensional Scaling*. 2nd Edition, Chapman & Hall/CRC, New York, 2001.
- D.L. DAVIES et D.W. BOULDIN : A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, 1979.
- P.T. DAVIES : *What is Evidence-Based Education ?* Jerry Lee Lecture presented to the 4th Annual Campbell Collaboration Colloquium, Washington D. C., 2004.
- J. DEUTSCH et J. SILBER : Measuring multidimensional poverty : An empirical comparison of various approaches. *Review of Income and Wealth*, 51(1):145–174, 2005.
- E. DIDAY : La méthode des nuées dynamiques. *Revue de Stat Appliquée*, 19 (2):19–34, 1971.
- E. DIDAY : Une représentation visuelle des classes empiétantes : les pyramides. *Revue RAIRO APII*, 20(5):475–526, 1986.
- E. DIDAY, J. LEMAIRE, J. POUGET et F. TESTU : *Éléments d'analyse des données*. Bordas, Paris, 1982.
- J.-Y. DUCLOS, D.E. SAHN et S.D. YOUNGER : Robust multidimensional poverty comparisons. *Economic Journal*, 116(514):943–968, 2006.
- R. ECHÁVARRI et I. PERMANYER : Ranking profiles of capability sets. *Social Choice and Welfare*, 31(3):521–535, 2008.
- D. ELLSBERG : Classic and current notions of measurable utility. *The Economic Journal*, 64(255):528–556, 1954.
- ENCARTA : *Encarta Poverty Definition*. Encarta World English Dictionary, North American Edition, 2010.
- M. ESTER, H.P. KRIEGEL, J. SANDER et X. XU : A density-based algorithm for discovering clusters in large spatial databases with noise. *In Second International Conference on Knowledge Discovery and Data Mining*, pages 226–231. AAAI Press, 1996.
- B.S. EVERITT : *Cluster Analysis*. Heineman Educational Books Ltd., 1980.
- G. FERRO-LUZZI, Y. FLÜCKIGER et S. WEBER : A cluster analysis of multi-dimensional poverty in Switzerland. CRAG HES-SO/HEG-GE/C–06/1/2–CH, Université de Genève, 2006.

- D. FILMER et L. PRITCHETT : Child mortality and public spending on health : How much does money matter ? Policy Research Working Paper Series 1864, The World Bank, 1997.
- D. FILMER et L. PRITCHETT : Estimating wealth effects without expenditure data or tears : An application to educational enrollments in states of India. *Demography*, 38(1):115–132, 2001.
- D.H. FISHER : Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, 2(2):139–172, 1987.
- E. FORGY : Cluster analysis of multivariate data : Efficiency versus interpretability of classification. *Biometrics*, 21(3):768–769, 1965.
- J. FOSTER, J. GREER et E. THORBECKE : A class of decomposable poverty measures. *Econometrica*, 52(3):761–66, 1984.
- J. FOSTER et A.F. SHORROCKS : Poverty orderings. *Econometrica*, 56(1):173–177, 1988a.
- J.E. FOSTER et A.F. SHORROCKS : Inequality and poverty orderings. *European Economic Review*, 32(2-3):654–661, 1988b.
- A. FUSCO : *La Contribution des Analyses Multidimensionnelles à la Compréhension et à la Mesure du Concept de Pauvreté : Application Empirique au Panel Communautaire des Ménages*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2005.
- C. GINI : Variability and mutability. *Journal of the Royal Statistical Society*, 76(6):619–622, 1913.
- A. GOMES DA SILVA : *Analyse des données évolutives : application aux données d'usage du Web*. These, Université Paris Dauphine - Paris IX, 2009.
- L.A. GOODMAN et W.H. KRUSKAL : Measures of association for cross classification. *Journal of the American Statistical Association*, 49:732–764, 1954.
- A.D. GORDON : *Classification*. 2nd Edition, Chapman & Hall/CRC, New York, 1999.
- J.C. GOWER : A general coefficient of similarity and some of its properties. *Biometrics*, 27(4):857–871, 1971.
- J.C. GOWER et B. ZIELMAN : Orthogonality and its approximation in the analysis of asymmetry. *Linear Algebra and its Applications*, 278(1-3):183–193, 1998.

- P.J.F GROENEN et M.V.D. VELDEN : Multidimensional scaling. Econometric Institute Report EI 2004-15, Erasmus University Rotterdam, 2004.
- S. GUHA, R. RASTOGI et K. SHIM : ROCK : A robust clustering algorithm for categorical attributes. *Information Systems*, 25(5):345–366, 2000.
- S. GUHA, R. RASTOGI et K. SHIM : CURE : An efficient clustering algorithm for large databases. *Information Systems*, 26(1):35–58, 2001.
- M. HALKIDI, Y. BATISTAKIS et M. VAZIRGIANNIS : On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 17(2-3):107–145, 2001.
- A.J. HARTIGAN : *Clustering Algorithms*. John Wiley & Sons, Inc, New York, USA, 1975.
- J.A. HARTIGAN et M.A. WONG : A k-means clustering algorithm. *Journal of the Royal Statistical Society : Applied Statistics*, 28(1):100–108, 1979.
- J. HAUGHTON et S.R. KHANDKER : *Handbook on Poverty and Inequality*. The International Bank for Reconstruction and Development/The World Bank, Washington, DC 20433, 2009.
- J.G. HIRSCHBERG, E. MAASOUMI et J.D. SLOTTJE : Cluster analysis for measuring welfare and quality of life across countries. *Journal of Econometrics*, 50 (1-2):131–150, 1991.
- H. HOTELLING : The generalization of Student's ratio. *The Annals of Mathematical Statistics*, 2(3):360–378, 1931.
- J.L. HUBERT et R.J. LEVIN : A general statistical framework for assessing categorical clustering in free recall. *Psychological Bulletin*, (83):1072–1080, 1976.
- L. HUBERT et P. ARABIE : Comparing partitions. *Journal of Classification*, (1):193–218.
- A.K. JAIN et C.R. DUBES : *Algorithms for Clustering*. Prentice-Hall, Inc, Englewood Cliffs, New Jersey 07632, 1988.
- A.K. JAIN, M.N. MURTY et P.J. FLYNN : Data clustering : a review. *ACM Computing Surveys*, 31:264–323, 1999.
- N. JARDINE et R. SIBSON : *Mathematical Taxonomy*. Wiley London, New York, 1971.
- J. KABUBO-MARIARA, A. WAMBUGU et S. MUSAU : Multidimensional poverty in Kenya : Analysis of maternal and child wellbeing. Cahiers de recherche PMMA 2011-12, PEP-PMMA, 2011.

- N. KAKWANI et J. SILBER : *Quantitative Approaches to Multidimensional Poverty Measurement*. Palgrave, Macmillan, 2008.
- V.Z. KANA, B. SOMÉ et A. TSOUKIÀS : A survey on multidimensional poverty measurement : a decision aiding perspective. Cahier du LAMSADE N° 313, LAMSADE, CNRS, 2011.
- V.Z. KANA, B. SOMÉ et A. TSOUKIÀS : A new methodology for multidimensional poverty measurement based on the capability approach. Cahier du LAMSADE N° 323, LAMSADE, CNRS, 2012.
- G. KARYPIS, E.-H. HAN et V. KUMAR : CHAMELEON : Hierarchical clustering using dynamic modeling. *IEEE Computer*, 32(8):68–75, 1999.
- L. KAUFMAN et P.J. ROUSSEEUW : *Finding Groups in Data : An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 2005.
- S. KLASSEN : Measuring poverty and deprivation in South Africa. *Review of Income and Wealth*, 46(1):33–58, 2000.
- T. KOHONEN : Neurocomputing : Foundations of research. chapitre Self-organized formation of topologically correct feature maps, pages 509–521. MIT Press, Cambridge, MA, USA, 1988.
- T. KOHONEN, M.R. SCHROEDER et T.S. HUANG : *Self-Organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2001.
- S. B. KOTSIANTIS : Supervised machine learning : A review of classification techniques. *Informatica*, 31:249–268, 2007.
- D.H. KRANTZ, R.D. LUCE, P. SUPPES et A. TVERSKY : *Foundation of Measurement : Additive and Polynomial Representations*. Academic Press, vol. 1, New York, 1971.
- J.B. KRUSKAL et M. WISH : *Multidimensional Scaling*. Sage Publications, Beverly Hills and London, 1978.
- J.W. KRZANOWSKI et T.Y. LAI'S : A criterion for determining the number of groups in a data set using Sum-of-Squares clustering. *Biometrics*, 44(1):23–34, 1988.
- L. LEBART, M. PIRON et A. MORINEAU : *Statistique exploratoire multidimensionnelle. Visualisation et inférence en fouilles de données*. Dunod, 2006.

- O. LELKES et E. ZÓLYOMI : *Poverty Across Europe : The Latest Evidence Using the EU-SILC Survey*. European Centre, 2008.
- C.A.K. LOVELL, S. RICHARDSON, P. TRAVERS et L. WOOD : Resources and functions : A new view of inequality in Australia. School of Economics Working Papers 1990-07, University of Adelaide, School of Economics, 1990.
- M.A. LUGO et E. MAASOUMI : Multidimensional poverty measures from an information theory perspective. Working Papers 85, ECINEQ, Society for the Study of Economic Inequality, 2008.
- E. MAASOUMI : The measurement and decomposition of multi-dimensional inequality. *Econometrica*, 54(4):991–97, 1986.
- E. MAASOUMI : A compendium to information theory in economics and econometrics. *Econometric Reviews*, 12(2):137–181, 1993.
- J. MACK et S. LANSLEY : *Poor Britain*. George Allen & Unwin, London, 1985.
- L. MARI : Epistemology of measurement. *Measurement, Elsevier Science Ltd*, 34 (1):17–30, 2003.
- J.D. MARSHALL : *The Old Poor Law, 1795-1834*. 2nd edition, Macmillan, London, 1985.
- K. MARX : *The Class Struggles in France, 1848 to 1850*. Published as a booklet by Engels in 1895. In Selected Works, Volume 1, Progress Publishers, Moscow (1969), 1850.
- K. MARX : *Capital : A Critical Analysis of Capitalist Production*. Volume 1. Sonnenschein, London, 1887.
- WORLD BANK : *World Development Report : Poverty*. DC : World Bank, Washington, 1990.
- WORLD BANK : *World Development Report : Attacking Poverty 2000/2001*. DC : World Bank, Washington, 2001.
- WORLD BANK : *Introduction to Poverty Analysis*. Poverty Manual, World Bank Institute, 2005.
- R.M. MCINTYRE et R.K. BLASHFIELD : A nearest-centroid technique for evaluating the minimum-variance clustering procedure. *Multivariate Behavioral Research*, 15(2):225–38, 1980.

- G.J. MCLACHLAN : *Discriminant Analysis and Statistical Pattern Recognition*. Wiley-Interscience, Wiley Series in Probability and Statistics, 2004.
- S. MILANO : *La pauvreté dans les pays riches - du constat à l'analyse*. Edition Nathan, Paris, 1992.
- J.S. MILL : *Principles of Political Economy*. Prometheus Books, London, 1848.
- J.S. MILL : *Utilitarianism*. 4th edition. Longmans, Green, Reader, and Dyer, 1871.
- G. MILLIGAN et M. COOPER : A study of standardization of variables in cluster analysis. *Journal of Classification*, 5(2):181–204, 1988.
- G.W. MILLIGAN : Clustering validation : results and implications for applied analyses. *Clustering and Classification : P.Arabie, L. J. Hubert, & G. De Soete (Eds)*, World Scientific Publ., River Edge, NJ, pages 341–375, 1996.
- G.W. MILLIGAN et M.C. COOPER : An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2):159–179, 1985.
- J. MORDUCH et M. SHARMA : Strengthening public safety nets from the bottom up. *Development Policy Review*, 20(5):569–588, 2002.
- A. MORINEAU : Note sur la caractérisation statistique d'une classe et les valeurs tests. *Bulletin Technique Centre Statistique Informatique Appliquées*, 2(1-2):20–27, 1984.
- J.P. NAKACHE et J. CONFAIS : *Approche Pragmatique de la Classification : Arbres Hiérarchiques, Partitionnements*. Editions Technip, 2004.
- V. NAVARRO : Development and quality of life : a critique of Amartya Sen's development as freedom. *International Journal of Health Services*, 30(4):661–74, 2000.
- R.T. NG et J. HAN : Efficient and effective clustering methods for spatial data mining. *In Proceedings of the 20th International Conference on Very Large Data Bases, VLDB '94*, pages 144–155, San Francisco, CA, USA, 1994. Morgan Kaufmann Publishers Inc.
- B. NOLAN et C.T. WHELAN : *Resources, Deprivation and Poverty*. Clarendon Press, Oxford, 1996.
- M.C. NUSSBAUM : Nature, function and capability : Aristotle on political distribution. Working Papers 31, UNU-WIDER, 1987.

- M.C. NUSSBAUM : *Women and Human Development : the Capabilities Approach*. Cambridge University Press, Cambridge, 2000.
- M.C. NUSSBAUM : Capabilities as fundamental entitlements : Sen and social justice. *Feminist Economics*, 9(2-3):33–59, 2003.
- M.C. NUSSBAUM : *Well-Being, Contracts and Capabilities*. In Lenore Manderson (ed.), *Rethinking Well-Being*, Perth : API Network, 2005.
- V. PARETO : *Cours d'Economie Politique professé à l'Université de Lausanne*. F. Rouge, Lausanne, 1896-97.
- V. PARETO : *Manual of Political Economy*. Macmillan, London, 1927.
- P. PERNY : Multicriteria filtering methods based on concordance and non-discordance principles. *Annals of Operations Research*, 80:137–167, 1998.
- P. PERNY et B. ROY : The use of fuzzy outranking relations in preference modeling. *Fuzzy Sets and Systems*, 49:33–53, 1992.
- A.C. PIGOU : *The Economics of Welfare*. 4th edition Macmillan, London, 1920.
- PNUD : *Rapport Mondial sur le Développement Humain*. Economica, Paris, 1990–2006.
- T. POGGE : Can the capability approach be justified? *Philosophical Topics*, 30 (2):167–228, 2002.
- G.W. PROTHERO : *Select statutes and other constitutional documents illustrative of the reigns of Elizabeth and James I*. 4th edition printed in 1934 by The Clarendon Press, Oxford, 1894.
- SAS PUBLISHING : *SAS Technical Report A-108, Cubic Clustering Criterion*. Sas Institute Inc., 1998.
- M. QIZILBASH : Vagueness and the measurement of poverty. Discussion Paper 2000-03, School of Economics and Social Studies, University of East Anglia, 2000.
- M. QIZILBASH : Vague language and precise measurement : the case of poverty. *Journal of Economic Methodology*, 10(1):41–58, 2003.
- R DEVELOPMENT CORE TEAM : *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011.

- R. RAM : Composite indices of physical quality of life, basic needs fulfilment, and income : A principal component representation. *Journal of Development Economics*, 11(2):227–247, 1982.
- W.M. RAND : *Journal of the American Statistical Association*, (336):846–850.
- M. RAVALLION : Comparaisons de la pauvreté : Concepts et méthodes. LSMS Document de Travail 122, Washington, DC : World Bank, 1996.
- M. RAVALLION : Poverty lines in theory and practice. LSMS Working Papers 133, Washington, DC : World Bank, 1998.
- M. RAVALLION : *Poverty Lines*. The New Palgrave Dictionary of Economics, 2nd edition, Larry Blume and Steven Durlauf (eds), Palgrave Macmillan, London, 2008.
- J. RAWLS : *A Theory of Justice*. Harvard University Press, 1971.
- C.J. Van RIJSBERGEN : *Information Retrieval*. Butterworth-Heinemann, 2nd édition, 1979.
- L. ROBBINS : Interpersonal comparisons of utility : a comment. *Economic Journal*, 48(192):635–641, 1938.
- F.S. ROBERTS : *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences*. Encyclopedia of Mathematics and its Applications (vol. 7), Addison-Wesley, 1979.
- I. ROBEYNS : *Capabilities and Theories of Justice*. In Enrica Chiappero Martinetti (eds), *Debating Global Society : Reach and Limits of the Capability Approach*, Milan : Feltrinelli, 2009.
- I. ROBEYNS : *Measuring Justice : Primary Goods and Capabilities*. Cambridge University Press, 2010.
- P. ROUSSEEUW : Silhouettes : a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(1):53–65, 1987.
- R.S. ROWNTREE : *Poverty*. MacMillan, London, 1901.
- R.S. ROWNTREE : *Poverty and the Welfare State*. Longmans Green, London, 1941.
- B. ROY : Classement et choix en présence de points de vue multiples : La méthode ELECTRE. *Revue Francaise d'Informatique et de Recherche Opérationnelle*, 8:57–75, 1968.

- B. ROY : ELECTRE III : Un algorithme de classements fondé sur une représentation floue des préférences en présence de critères multiples. *Cahiers du CERO*, 20(1):3–24, 1978.
- B. ROY : *Méthodologie multicritère d'aide à la décision*. Economica, Collection Gestion, Paris, 1985.
- B. ROY : *Multicriteria Methodology for Decision Aiding*. Kluwer Academic, Dordrecht, 1996.
- B. ROY et D. BOUYSSOU : *Aide Multicritère à la Décision : Méthode et Cas*. Economica, Collection Gestion, Paris, 1993.
- W.G. RUNCIMAN : *Relative Deprivation and Social Justice*. Routledge and Kegan Paul, London, 1966.
- D.E. SAHN et D. STIFEL : Exploring alternative measures of welfare in the absence of expenditure data. *Review of Income and Wealth*, 49(4):463–489, 2003.
- D.E. SAHN et D.C. STIFEL : Poverty comparisons over time and across countries in Africa. *World Development*, 28(12):2123–2155, 2000.
- R. SAITH : Capabilities : the concept and its operationalisation. QEH Working Papers QEHWPS-66, Queen Elizabeth House, University of Oxford, 2001.
- B. SAMI : Measuring poverty in a multidimensional perspective : a review of literature. Cahiers de recherche PMMA 2005-07, PEP-PMMA, 2005.
- B. SAMI et A.-R. EL LAHGA : Les mesures multidimensionnelles de la pauvreté : une application sur l'Afrique du Sud et l'Égypte. Working Papers 06-39, CIR-PÉE, 2006.
- B. SAMI et A.-R. EL LAHGA : Comparaisons ordinales robustes de la pauvreté multidimensionnelle : Afrique du Sud et Egypte. *Revue d'Economie du Développement*, 22(1):5–36, 2008.
- J. SANDER, M. ESTER, H.-P. KRIEGEL et X. XU : Density-based clustering in spatial databases : The algorithm GDBSCAN and its applications. 2(2):169–194, 1998.
- B. SANDERSON : *Greece & Rome to 30 BC*. Ethics of civilization, volume 4, World Peace Communications, 2004.
- G. SAPORTA : *Probabilités, Analyse des Données et Statistique*. Éditions TECHNIP, Paris, 2006.

- D. SEERS : The meaning of economic development. *International Development Review*, 11(4):9–30, 1969.
- A.K. SEN : *Collective Choice and Social Welfare*. San Fransisco : Holden-Day, 1970.
- A.K. SEN : Poverty : An ordinal approach to measurement. *Econometrica*, 44 (2):219–31, 1976.
- A.K. SEN : Equality of what? The Tanner lecture on human values, Stanford University, 1979a.
- A.K. SEN : Issues in the measurement of poverty. *Scandinavian Journal of Economics*, 81(2):285–307, 1979b.
- A.K. SEN : Utilitarianism and welfarism. *Journal of Philosophy*, 76(9):463–489, 1979c.
- A.K. SEN : *Commodities and Capabilities*. Elsevier Science Publishers B.V., North-Holland, 1985.
- A.K. SEN : *Inequality reexamined*. Russell Sage Foundation, New York, 1992.
- A.K. SEN : *Capability and Well-being*. in Nussbaum M. and Sen A.(eds), *The Quality of Life*, Oxford University Press, 1993.
- A.K. SEN : *Development as Freedom*. Alfred A Knopf, 1999.
- A.K. SEN : Amartya Sen and the thousand faces of poverty. Interview of Sen by Quesada Charo, Inter-American Development Bank, 2001.
- A. SHAH : Structural adjustment – a major cause of poverty. Global issues, <http://www.globalissues.org/article/3/>, 2010.
- C.E. SHANNON : A mathematical theory of communication. *Bell system technical journal*, 27:379–423, 1948.
- H.S. SIDGWICK : *The Methods of Ethics*. Hackett Publishing Company, London, 1874.
- J. SILBER : Measuring poverty : Taking a multidimensional perspective. *Hacienda Pública Española – Revista de Economía Pública*, 182(3):29–73, 2007.
- P. SLACK : *The English Poor Law, 1531-1782*. Macmillan, London, 1990.
- D. SMITH : *Poverty Rate hits 15-year High*. REUTERS September 17 edition, Washington, 2010.

- S.P. SMITH et A.K. JAIN : Testing for uniformity in multidimensional data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(1):73–81, 1984.
- R.R. SOKAL et C.D. MICHENER : A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38:1409–1438, 1958.
- T. SØRENSEN : A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab*, 5(4):1–34, 1957.
- S.S. STEVENS : On the theory of scales of measurement. *American Association for the Advancement of Science*, 103(2684):677–680, 1946.
- G. STIGLER : The development of utility theory I. *Journal of Political Economy*, 58(4):307–327, 1950.
- P. SUPPES : *Measurement empirical meaningfulness and three-valued logic*. in Churchman and Ratoosh (eds.), *Measurement : Definition and Theories*, John Wiley & Sons, New York, 1959.
- R. SUZUKI et H. SHIMODAIRA : Pvcust : an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics (Oxford, England)*, 22(12):1540–1542, 2006.
- H. THEIL : *Economics and Information Theory*. American Elsevier Publishing Company, New York, 1967.
- R. TIBSHIRANI, G. WALTHER et T. HASTIE : Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society Series B*, 63(2):411–423, 2001.
- S.W. TORGERSON : *Theory and methods of scaling*. Wiley, New York, 1967.
- P. TOWNSEND : *Economics and Information Theory*. University of California Press, California, 1979.
- P. TOWNSEND : Deprivation. *Journal of Social Policy, Cambridge University Press*, 16, 1987.
- A. TSOUKIÀS : On the concept of decision aiding process : an operational perspective. *Annals of Operations Research*, 154(1):3–27, 2007.
- A. TSOUKIÀS : From decision theory to decision aiding methodology. *European Journal of Operational Research*, 187(1):138–161, 2008.

- A. TSOUKIÀS, P. PERNY et P. VINCKE : From concordance-discordance to the modelling of positive and negative reasons in decision aiding. *In Aiding Decisions with Multiple Criteria : Essays in Honour of Bernard Roy*, pages 147–175. Kluwer Academic Publishers, 2002.
- K.-Y. TSUI : Multidimensional generalizations of the relative and absolute inequality indices : The Atkinson-Kolm-Sen approach. *Journal of Economic Theory*, 67(1):251–265, 1995.
- K.-Y. TSUI : Multidimensional poverty indices. *Social Choice and Welfare*, 19(1):69–93, 2002.
- S. TUFFÉRY : *Data Mining et Statistique Décisionnelle : l'Intelligence des Données*. Éditions Technip, Paris, 2010.
- UNDP : Indicateurs pour le suivi des Objectifs du Millénaire pour le Développement - Définitions, Raison d'être, Concepts et Sources de données. United nations development report, UNDP, 2003.
- UNDP : The real wealth of nations : pathways to human development. Human development report, UNDP, 2010.
- B.M.S. VAN PRAAG : *Individual Welfare Functions and Consumer Behaviour*. North Hollande Publishing Company, Amsterdam, 1968.
- Ph. VINCKE : *Multicriteria Decision-Aid*. J. Wiley, New York, 1992.
- W. WANG, J. YANG et R.R. MUNTZ : STING : A Statistical Information Grid Approach to Spatial Data Mining. *In Twenty-Third International Conference on Very Large Data Bases*.
- J.H. WARD : Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.
- K. WATKINS : *The Oxfam Poverty Report*. Oxfam United Kingdom, Oxford, 1995.
- H. WATTS : *An Economic Definition of Poverty*. in D. P. Moynihan (Ed.), *On Understanding Poverty*, Basic Books, New York, 1968.
- S.M. WEISS et C.A. KULIKOWSKI : *Computer Systems That Learn : Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning and Expert Systems*. Morgan Kaufmann, San Francisco, CA, 1991.
- K.-G. WOO, J.-H. LEE, M.-H. KIM et Y.-J. LEE : Findit : a fast and intelligent subspace clustering algorithm using dimension voting. *Information & Software Technology*, 46(4):255–271, 2004.

- Y. XU : Functioning, capability and the standard of living – an axiomatic approach. *Economic Theory*, 20(2):387–399, 2002.
- M.B. YÉLÉ : Dominance stochastique et pauvreté multidimensionnelle dans les pays de l’UEMOA. Working papers, CIRPÉE, 2007.
- M.B. YÉLÉ : Multidimensional poverty dominance : Statistical inference and an application to West Africa. Working papers, CIRPÉE, 2008.
- A.F. YOUNG et E.T. ASHTON : *British Social Work in the Nineteenth Century*. Routledge, London, 1956.
- L.A. ZADEH : Fuzzy sets. *Information and Control*, 8 :338-53, 1965.
- J. ZATTLER : The effects of structural adjustment programmes. *Intereconomics, Development Strategy*, 24(6):282–289, 1989.
- T. ZHANG, R. RAMAKRISHNAN et M. LIVNY : BIRCH : A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery*, 1(2): 141–182, 1997.
- B. ZHENG : Aggregate poverty measures. *Journal of Economic Surveys*, 11(2): 123–162, 1997.
- B. ZHENG : On the power of poverty orderings. *Social Choice and Welfare*, 16 (3):349–371, 1999.
- B. ZHENG : Poverty orderings : A graphical illustration. *Social Choice and Welfare*, 18(1):165–178, 2001.

Vu :
Le Président
M.

Vu :
Les Suffrageants
M.

Vu et permis d'imprimer :
Le Vice-Président du Conseil Scientifique Chargé de la Recherche de l'Université
Paris Dauphine

Construction d'indices : mesurage multidimensionnel de la pauvreté

Résumé :

L'émergence rapide et incontrôlée d'une volonté d'éradication de la pauvreté dans les pays sous-développés, les pays en développement et les pays développés, sont à l'origine de nouvelles exigences en termes de méthodologie de lutte contre la pauvreté. Or la prise en compte du concept de développement durable requiert une analyse holistique de certaines politiques de réduction de la pauvreté. Cela nécessite une échelle d'étude multidimensionnelle plus large que celle du critère unique. Ce travail concerne la mise en place d'une méthode d'analyse multi-indicateurs basée sur l'utilisation des techniques de clustering dérivées de l'analyse des données et l'aide à la décision. L'hypothèse principale est qu'il n'existe pas de mesure "objective" de la pauvreté (un phénomène social particulièrement complexe), mais seulement une mesure opérationnelle destinée à une certaine politique de réduction de la pauvreté dans une région donnée.

La question de la signifiante est ainsi analysée à la fois d'un point de vue théorique (théorie du mesurage) et d'un point de vue opérationnel (élaboration et implémentation des politiques appropriées). Cette mesure multidimensionnelle de la pauvreté, signifiante pour une certaine politique de réduction de la pauvreté et pour un certain cadre décisionnel observé, est validée par une simulation appliquée aux bases de données sur la pauvreté au Burkina Faso.

Mots clés : Clustering, aide à la décision, pauvreté, cluster diagnostic, mesure signifiante, approche par les capacités, élaboration des politiques, affectation floue.

Construction of Indexes : Multidimensional Poverty Measurement

Abstract :

The rapid and uncontrolled emergence of a commitment to poverty eradication in under-developed countries, developing countries and developed countries, are causing new requirements in terms of methodology for fighting against poverty. Taking into account the concept of sustainable development requires a holistic analysis of certain poverty reduction policies. This requires a broader multidimensional study which is better than a single criterion approach.

This work concerns the development of a method for analysing multi-indicators based on the use of clustering techniques derived from data analysis and decision aid methods. The main hypothesis is there is no “objective” poverty measurement (a social phenomenon particularly complex), but only an operational measurement for a certain poverty reduction policy in a given region.

The issue of meaningfulness is thus analyzed both from a theoretical point of view (measurement theory) and from an operational one (policy effectiveness). This multidimensional poverty measurement is meaningful for some poverty reduction policy and for a certain decisional framework. We have validated our methodology by a simulation applied to databases on poverty from Burkina Faso.

Keywords : Clustering, decision aiding, poverty, cluster diagnosis, meaningful measurement, capabilities approach, policy making, fuzzy assignment.
