

Ministère de l'Enseignement Supérieur de la  
Recherche Scientifique et de l'Innovation

Secrétariat Général

Université Polytechnique de Bobo Dioulasso

Unité de Formation et de Recherche en  
Sciences et Techniques

Licence en Statistiques Informatique

Ministère de la Santé

Secrétariat Général

Centre MURAZ

Département de Recherche Clinique

Equipe d'Appui Méthodologique et de  
Formation



Rapport de stage pour l'obtention de la Licence en Statistiques-Informatique

Sous le thème:

# VALIDATION D'UN PLAN DE SONDAGE A BASE AREOLAIRE DANS LA COMMUNE URBAINE DE BOBO DIOULASSO

Elaboré par

Do Edmond SANOU

Encadré par

**Directeur de rapport**

Dr Boureima SANGARE

Enseignant-Chercheur

Université Polytechnique de Bobo Dioulasso

**Maître de stage**

Dr Serge SOMDA

Méthodologiste-Biostatisticien

Equipe Appui Méthodologique et Formation

Centre MURAZ

## Dédicace

A Celui qui est, qui était, et qui vient.

A mon père et à ma mère, mon héros et mon héroïne, pour les sacrifices consentis en vue de m'élever dignement et d'assurer mon éducation dans les meilleures conditions.

A mon frère et à ma sœur, mes appuis, avec qui j'ai grandi dans le même nid, en témoignage de mon affection fraternelle et de ma profonde tendresse.

A mes amis de toujours, en souvenir de notre sincère et profonde amitié et des moments agréables que nous avons passés ensemble.

Toutes les lettres ne sauraient trouver les mots qu'il faut.

Tous les mots ne sauraient exprimer la gratitude,

L'amour, le respect et la reconnaissance.

Aussi, c'est tout simplement que

Je vous dédie ce travail

SANOU Do Edmond

## Avant-propos

Le recueil d'informations fiables, leur gestion informatique dans des bases de données, l'analyse statistique qui en fournit la synthèse et la communication des résultats sont des éléments indispensables d'aide à la décision dans tous les secteurs d'activité.

La formation de Licence en Statistiques-Informatique de l'Université Polytechnique de Bobo Dioulasso vise plus particulièrement à pallier la carence actuelle de compétences au confluent de la statistique et de l'informatique avec l'économie. C'est un secteur où la demande en personnels qualifiés est de plus en plus marquée dans nos pays en voie de développement.

Les diplômés de la Licence en Statistiques-Informatique acquièrent ainsi une triple compétence en statistique, informatique et en économie. Ils ont une démarche assez critique et réfléchie devant des données, sachant identifier celles qui sont pertinentes et les analyser avec les principaux outils de la Statistique.

- Comprendre les problématiques et les enjeux d'un secteur d'activité,
- faire preuve d'autonomie en programmation, gestion des bases de données et utilisation de l'informatique communicante
- assurer la qualité des données
- savoir utiliser les logiciels spécialisés de gestion des données et d'analyse statistique
- avoir une capacité d'expression et de communication adaptée aux nécessités du travail en équipe

... sont entre autres les principaux objectifs de cette licence.

La scolarité se déroule en trois années à l'issue desquelles les étudiants, disposant d'un socle de connaissances théoriques réalisent un stage d'une durée minimale de 12 semaines. Ce stage est évalué sur la base d'un rapport et d'une soutenance.

C'est dans ce cadre que nous avons été admis au sein de l'Equipe d'Appui Méthodologique et de Formation du Centre MURAZ pour notre stage de fin de cycle allant du 15 janvier 2016 au 2 septembre 2016 (30 semaines) portant sur un projet de recherche visant à mettre en place un système de sondage à base aréolaire dans la commune urbaine de Bobo Dioulasso. Le stage s'est déroulé en alternance du 15 janvier au 2 mai 2016 (14 semaines) puis à temps plein du 2 mai au 2 septembre 2016 (16 semaines). Le projet traitant d'un sujet assez complexe « Développement et validation d'un plan de sondage à base aréolaire dans la commune urbaine de Bobo Dioulasso », nous avons été deux étudiants à travailler là-dessus. Dans le souci de présenter un travail individuel en vue de l'évaluation en soutenance, le présent rapport traitera du thème suivant : « Validation d'un plan de sondage à base aréolaire dans la commune urbaine de Bobo Dioulasso ». Il s'articule autour des hypothèses et résultats du rapport Statistique Spatiale : Cartographie et modélisation spatiale par krigeage d'Armel SOUBEIGA, notre co-stagiaire.

# 1. Introduction

Les principales données démographiques au Burkina Faso proviennent de l'Institut National de la Statistique et de la Démographie (INSD) et sont disponibles en 2016 à l'échelle de régions, d'arrondissements, de communes ou de Zones de Dénombrement (ZD). Or de plus en plus d'études nécessitent une connaissance précise ou approximée du nombre de ménages ou d'individus au sein de périmètres spécifiques qui ne font pas obligatoirement partie des périmètres techniques ou administratifs de l'INSD. L'absence de "base de sondage liste"<sup>1</sup> et de données récentes rendent souvent laborieux l'élaboration de plans de sondage fiables et rigoureux (probabilistes). Il devient de plus en plus impératif de connaître finement les poids de population impactés par tel ou tel phénomène pour des raisons de santé publique, économique ou administrative. Un projet de recherche réalisé pour le Centre MURAZ par l'Equipe d'Appui Méthodologique et Formation (EAMF) durant le premier semestre de 2016 a permis de mettre en place un système de sondage à base aréolaire. Cette méthode de sondage apparaît aisément reproductible et repose sur des outils facilement mobilisables. Cependant la question de sa validité reste posée. La validation d'une méthode est la procédure par laquelle on démontre, preuves expérimentales à l'appui, que les performances de la méthode permettent de répondre aux exigences de l'usage auquel elle est destinée [1]. L'objectif de la présente étude est donc de proposer une procédure d'échantillonnage, qui s'appuie sur des estimations de population en des zones géographiques de la commune urbaine de Bobo Dioulasso. Elle s'articule autour de trois principaux points. D'abord, nous présentons de façon sommaire ladite méthode d'échantillonnage, ses caractéristiques, ses estimations et des cartographies de la ville de Bobo Dioulasso. Ensuite, nous abordons les techniques de validation. Enfin, nous proposons une application du plan de sondage.

## 1.1. Contexte et justification du thème de recherche

### 1.1. Contexte géographique, démographique

La province du Houet créée le 15 Septembre 1983, est l'une des 45 provinces que compte le Burkina Faso. Située à l'Ouest, elle couvre une superficie de 11540 km<sup>2</sup> soit 4,21% du pays [2]. Elle s'étend sur trois bassins versants internationaux qui sont la Comoé, la Volta et le Niger. Elle fait partie des provinces les plus arrosées du Burkina.

Son chef-lieu Bobo-Dioulasso, est la deuxième ville du pays et se situe à 365 km de la capitale Ouagadougou. De superficie 136,78 km<sup>2</sup>, la commune de Bobo Dioulasso comprend une zone urbaine et une zone périurbaine ou rurale. Elle est organisée en 7 arrondissements regroupant 33 secteurs selon le nouveau découpage communal de 2012. Au recensement de 2006, la ville comptait 489 967 habitants avec une densité de l'ordre de 3 582 habitants au km<sup>2</sup> [3].

### 1.2. Justification

---

<sup>1</sup> Liste exhaustive et mise à jour de tous les individus de la population sans omission ni double décompte.

Conscient des problèmes de santé de la population qui ne cesse de croître, il est nécessaire d'avoir des connaissances sur certaines caractéristiques de cette population (base d'informations) afin de mieux définir les politiques de santé. L'absence de base de sondage et de données récentes rend souvent laborieux l'élaboration de plans de sondage fiables et rigoureux (probabilistes). Plusieurs méthodologies d'approches existent pour échantillonner les ménages. Cependant deux d'entre elles ont particulièrement retenues notre attention car faisant partie des plus utilisées.

### **1.2.1. La méthode de l'Institut National de la Statistique et de la Démographie**

C'est l'organe officiel de l'Etat en matière d'information statistique. L'INSD réalise souvent des enquêtes dans le cadre du suivi des conditions de vie des ménages. Il utilise le plus souvent un plan de sondage à deux degrés.

- Au premier degré, sont tirées les unités primaires (UP) encore appelées ZD.
- Au second degré, sont tirées les unités secondaires (US) avec une probabilité égale pour toutes les US de l'UP.

En effet, une opération de dénombrement est menée au préalable pour énumérer tous les US des UP sélectionnées afin de constituer une base de sondage à ce niveau, d'où l'appellation ZD attribuée aux UP. La base ainsi conçue servira pour le tirage des unités secondaires et la mise à jour de la taille des ZD pour l'estimation des coefficients d'extrapolation. L'échantillon sélectionné est donc probabiliste à deux degrés.

### **1.2.2. La méthode classique des quotas et itinéraires**

La méthode des quotas est une des méthodes d'échantillonnage non probabilistes les plus communes. L'échantillonnage est effectué jusqu'à ce qu'un nombre déterminé d'unités (quotas) soient sélectionnées dans diverses sous populations. On commence tout d'abord, par stratifier l'univers, mais au lieu de tirer au hasard et à l'intérieur des strates les éléments qui vont constituer l'échantillon, les enquêteurs sont laissés libres d'interroger tout sujet qu'ils peuvent rencontrer pourvu que la condition  $x$  pour cent de chaque strate soit remplie,  $x$  étant la taille de l'échantillon exprimée en pourcentage par rapport à celle de l'univers. Par exemple, si la population comprend 100 enfants et 50 adultes et qu'il faille tirer un échantillon de taille 30, 20 enfants et 10 adultes seront interviewés. L'échantillonnage par quotas peut être considéré préférable à d'autres formes d'échantillonnage non-probabilistes parce qu'il faut inclure des membres de sous-population différentes.

La méthode des itinéraires est également une méthode non probabiliste. Elle consiste à imposer à l'enquêteur un itinéraire en lui indiquant exactement les points du circuit où il doit procéder à une interview. L'identification de ces points d'enquêtes résulte de la combinaison de tirages aléatoires parmi des coordonnées géographiques et de la consultation d'une carte détaillée. L'itinéraire peut également être repris sur un schéma. L'enquêteur reçoit une carte qui mentionne le nom de la commune où les enquêtes auront lieu, l'adresse (rue et numéro) à laquelle la première enquête doit être effectuée. L'enquêteur poursuivra ensuite son chemin suivant le plan, en ayant soin de respecter les consignes qui lui auront été données.

### 1.2.3. Comparaison des méthodes

La première approche est complexe, demande davantage de temps (car se déroulant en 2 phases) et coûte habituellement plus cher que l'échantillonnage non probabiliste. Étant donné que les unités de la population sont sélectionnées au hasard, et que la probabilité d'inclusion de chaque unité peut être calculée, il est cependant possible de faire des estimations fiables, ainsi que des estimations de l'erreur d'échantillonnage, et de formuler des inférences au sujet de la population.

La méthode classique des quotas et itinéraires est un moyen de sélectionner des unités d'une population à l'aide d'une méthode subjective. Il n'est pas nécessaire d'avoir une base de sondage complète pour ce type de sondage. C'est donc un moyen rapide, facile et bon marché d'obtenir des données. Cependant strictement parlant, elle ne permet pas l'inférence statistique et est moins fiable. En effet, il n'est pas évident qu'il soit possible de généraliser et d'appliquer les résultats de l'échantillon à toute la population en ce sens que les résultats sont plus variables d'un échantillon à l'autre.

## 1.2. Le thème de recherche

Le thème de recherche s'intitule « Développement et validation d'un plan de sondage à base aréolaire dans la commune urbaine de Bobo Dioulasso ».

### 1.2.1. Objectif général

Notre projet avait pour objectif principal la mise en place d'un système de sondage à base aréolaire qui allie démarche scientifique valide et facilité de mise en œuvre à moindre coût dans la commune urbaine de Bobo Dioulasso. Une base aréolaire est une base de sondage où l'unité d'échantillonnage est une fraction du territoire occupé par la population étudiée.

Le système de sondage établi propose une démarche simple et pratique qui pourra servir à toute institution ou à tout particulier voulant mener un sondage dans ladite commune. Aussi, permettra-t-il de renforcer les capacités du Centre MURAZ à travers son Equipe d'Appui Méthodologique et de Formation (EAMF) dans la collecte de l'information et de la recherche médicale.

### 1.2.2. Objectifs spécifiques

De façon plus spécifique, la présente recherche vise à :

- identifier des lieux-dits<sup>4</sup>,
- préparer une cartographie de la ville de Bobo-Dioulasso présentant les lieux-dits,
- estimer la densité de population au niveau des lieux-dits sélectionnés,
- établir une base de sondage aréolaire de la commune urbaine de Bobo Dioulasso,
- décrire le plan de sondage utilisant la base aréolaire,
- définir un estimateur du total associé au plan de sondage,
- définir la variance associée à l'estimateur,

---

<sup>4</sup> Un lieu-dit est un endroit remarquable, de faible étendue auquel est associé un nom propre.

- appliquer le plan de sondage à un cas pratique.

## 1.3. Le thème de stage

Nous avons été deux étudiants à travailler sur le thème de recherche. Dans le souci de présenter un travail individuel pour l'évaluation de la soutenance, deux thèmes de stage en ont découlé :

- « Statistique Spatiale : Cartographie et modélisation spatiale par krigeage » : ce thème a fait l'objet du rapport d'Armel SOUBEIGA. Il a permis de mettre en place une base aréolaire de la commune urbaine de Bobo Dioulasso.
- « Validation d'un plan de sondage à base aréolaire dans la commune urbaine de Bobo Dioulasso » : ce thème est l'objet du présent rapport de stage.

### 1.3.1. Objectifs

La présente étude vise à :

- décrire le plan de sondage utilisant la base aréolaire,
- définir un estimateur du total associé au plan de sondage,
- définir la variance associée à l'estimateur,
- appliquer le plan de sondage.

### 1.3.2. Hypothèse

Cette étude repose sur l'hypothèse principale que la base de sondage aréolaire établie est exhaustive, récente et fiable.

## 1.4. Présentation de la structure d'accueil

Le Centre MURAZ est l'institution la plus ancienne dans le paysage burkinabè de la recherche scientifique. Fondé en 1939 par les colons français, il était à l'origine, un bureau de contrôle des maladies endémiques majeures, y compris la maladie du sommeil, l'onchocercose, le paludisme et les maladies à potentiel épidémique. Depuis 2001, c'est un institut public de recherche pour la santé avec pour mission de contribuer à la prévention, au diagnostic et au contrôle des maladies transmissibles et non-transmissibles à travers la provision de résultats de qualité en recherche biomédicale, la formation des chercheurs et des agents de santé et l'expertise technique en biologie clinique, en sciences humaines et en santé publique.

Le Centre MURAZ a trois grands atouts dans le domaine de la recherche : une plateforme élargie d'outils de biologie expérimentale et clinique, des équipes multidisciplinaires de recherche, un riche réseau de partenaires techniques et financiers au niveau national, régional et international. La plateforme de laboratoires couvre l'entomologie, la bactériologie, la mycobactériologie, la virologie, l'immunologie, l'hématologie, la biochimie et la biologie moléculaire. En plus, le

centre MURAZ dispose d'une Equipe d'appui Méthodologique et de Formation pour la saisie, la validation et les analyses statistiques. Nous avons effectué notre stage au sein de cette équipe.

L'Equipe d'Appui Méthodologique et de Formation est une unité entièrement consacrée à l'amélioration de la qualité de la recherche scientifique. Elle a pour ambitions :

- d'assurer une masse critique d'expertise méthodologique hautement spécialisée et multidisciplinaire aux chercheurs du Centre MURAZ et de ses institutions partenaires ;
- de développer la formation dans les disciplines contributives à la méthodologie de la recherche ;
- d'aider les chercheurs à concevoir des projets de recherche pertinents, à effectuer des analyses biostatistiques, à gérer les données, à valoriser la diffusion des résultats, et à veiller à ce que les études respectent toutes les normes réglementaires pertinentes ;
- de répondre aux questions et aux besoins des utilisateurs de connaissances en matière de recherche en proposant à ces derniers des interventions axées sur les solutions et en les aidant à appliquer ces solutions.

## 2. Matériel et méthodes

Dans cette partie, nous rappelons les principaux outils de la théorie des sondages et nous introduisons des notations qui nous seront utiles dans la suite du document.

### 2.1. Notations

#### 2.1.1. Description de la population

On appelle population l'ensemble des éléments que l'on étudie : les éléments qui la composent sont appelés individus. Nous souhaitons obtenir de l'information concernant une population  $U$ . Cette population est supposée contenir un nombre fini d'individus  $N$  appelé *taille* de la population. Ces individus sont repérés par un indice  $i = 1, \dots, N$ . On note :

$$U = \{U_1, \dots, U_N\} \equiv \{1, \dots, N\}$$

En assimilant un individu  $U_i$  à son label  $i$ .

Soit  $Y$  une variable qui peut être mesurée sur chacun des individus de  $U$  mais dont la valeur sur chacun de ces individus est inconnue.  $Y$  est appelée variable d'intérêt : la valeur prise par  $Y$  sur le  $i^{\text{ème}}$  individu est notée  $Y_i$ .

Seront définis comme suit :

|   |   |
|---|---|
| ▪ Le total des valeurs prises par $Y$ sur $U$ : | $T = \sum_{i=1}^N Y_i = N\bar{Y}$                       |
| ▪ La moyenne :                                  | $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$                |
| ▪ La variance :                                 | $\sigma^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2$ |

#### 2.1.2. Description de l'échantillon

Si l'on enquêtait l'ensemble de la population (recensement) sur la/les variable(s) d'intérêt, il serait théoriquement possible d'obtenir des estimations exactes pour les différentes fonctionnelles. On se heurte à des difficultés pratiques à savoir les moyens financiers et humains à mettre en œuvre.

On se contente donc généralement d'enquêter sur une partie des individus de la population appelée échantillon. Nous supposons ici que l'échantillon est sélectionné au moyen d'un plan de sondage  $p$  sans remise, i.e d'une loi de probabilité sur l'ensemble des parties de  $U$ .  $p$  vérifie donc

$$\forall s \subset U \quad p(s) \geq 0$$

avec

$$\sum_{s \subset U} p(s) = 1$$

Nous notons  $S$  l'échantillon aléatoire de taille  $n$ .

Seront définis comme suit, sur  $S$  :

|                           |   |
|---------------------------|---|
| ▪ la moyenne empirique :  | $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$                      |
| ▪ le total des valeurs :  | $t = \sum_{i=1}^n y_i$  |
| ▪ la variance empirique : | $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$ |

### 2.1.3. Les probabilités d'inclusion

Nous notons  $I_i$  l'indicatrice d'appartenance à l'échantillon pour l'unité  $i$  de  $U$ .

Pour un plan de sondage  $p$ , on appellera probabilité d'inclusion d'ordre 1 de l'unité  $i$  la probabilité  $\pi_i$  qu'a cette unité d'être retenue dans l'échantillon. Cette probabilité dépend du plan de sondage.

On appelle probabilité d'inclusion d'ordre 2 la probabilité  $\pi_{ij}$  que deux unités distinctes  $i$  et  $j$  soient retenues conjointement dans l'échantillon.

Pour toutes les unités  $i, j \in U$ , les variables indicatrices vérifient les propriétés suivantes :

$$E(I_i) = \pi_i$$

$$V(I_i) = \pi_i(1 - \pi_i)$$

$$Cov(I_i, I_j) = \pi_{ij} - \pi_i\pi_j$$

## 2.2. Estimation par les valeurs dilatées

L'objet d'un sondage est d'estimer une fonction de la variable d'intérêt

$$\theta = \theta(Y_i), i \in U$$

que l'on appellera paramètre d'intérêt ou fonctionnelle, et d'évaluer la précision de cette estimation (sous forme de variance). Parmi les fonctionnelles, on trouve le total de la variable, la moyenne, la variance et la dispersion (ou variance corrigée). Nous nous intéresserons dans le cadre de cette étude à l'estimation du total. Cela implique que

$$\theta(Y_i) = T(Y_i)$$

## Remerciements

Tout d'abord, nous tenons à remercier, le corps professoral de l'Unité de Formation et de Recherche en Sciences et Techniques de l'Université Polytechnique de Bobo Dioulasso ainsi que le coordonnateur de la Licence en Statistiques-Informatique, le Pr Adama OUEDRAOGO, pour avoir assuré la partie théorique de notre formation. Nous remercions vivement, notre directeur de rapport, le Dr Boureima SANGARE, pour les critiques et conseils qu'il nous a apportés concernant les missions évoquées dans ce rapport.

Ensuite, nous adressons nos remerciements au Dr Serge SOMDA, chef de l'Equipe d'Appui Méthodologique et de Formation du Centre MURAZ, qui a bien voulu nous confier ce projet de recherche en totale adéquation avec nos attentes. En sa qualité de maître de stage, nous le remercions également pour le temps passé ensemble et le partage de son expertise au quotidien. Nous remercions le personnel et les stagiaires de l'Equipe d'Appui Méthodologique et de Formation pour leur accueil, leur esprit d'équipe, en particulier Armel SOUBEIGA avec qui nous avons réalisé ce projet de recherche.

Enfin, nous tenons à remercier toutes les personnes qui nous ont conseillé et relu lors de la rédaction de ce rapport de stage notamment M Guillaume Chauvet, enseignant chercheur à l'Ecole Nationale de la Statistique et de l'Analyse de l'Information, M Arsène ZONGO, Assistant en suivi-évaluation à Counterpart Burkina et Yacouba KONE, notre camarade de promotion.

Si pour toute unité  $i$  de  $U$  on a  $\pi_i > 0$ , alors

$$\hat{T}_{HT} = \sum_{i=1}^n \frac{Y_i}{\pi_i}$$

est un estimateur sans biais de  $T$ .

$\hat{T}_{HT}$  est appelé estimateur de Horvitz-Thompson (ou  $\pi$ -estimateur) du total  $T$ . On utilise également la dénomination d'estimateur par valeurs dilatées, car il s'agit d'un estimateur qui affecte un poids  $d_i = 1/\pi_i$  à chaque unité de l'échantillon. On dit encore que l'unité  $i$  de l'échantillon représente  $1/\pi_i$  unités de la population dans l'estimation du total.

L'estimateur de Horvitz-Thompson est le seul estimateur linéaire sans biais d'un total de la forme

$$\sum_{i \in U} w_i 1_{i \in S} Y_i$$

où les  $w_i$  soient indépendants de l'échantillon. Ici  $w_i = 1/\pi_i$  est égal à l'inverse de la probabilité d'inclusion, c'est-à-dire au poids d'échantillonnage qui est disponible dans la base de sondage.

### 2.2.1. Précision et précision estimée de l'estimateur

Le  $\pi$ -estimateur de total a pour variance :

$$V(\hat{T}_{HT}) = \sum_{i=1}^N \sum_{j=1}^N \frac{Y_i Y_j}{\pi_i \pi_j} (\pi_{ij} - \pi_i \pi_j)$$

Cette variance peut être estimée sans biais par :

$$\hat{V}(\hat{T}_{HT}) = \sum_{i=1}^n \frac{Y_i^2}{\pi_i^2} (1 - \pi_i) + \sum_{i=1}^n \sum_{j=1}^n \frac{Y_i Y_j}{\pi_i \pi_j \pi_{ij}} (\pi_{ij} - \pi_i \pi_j)$$

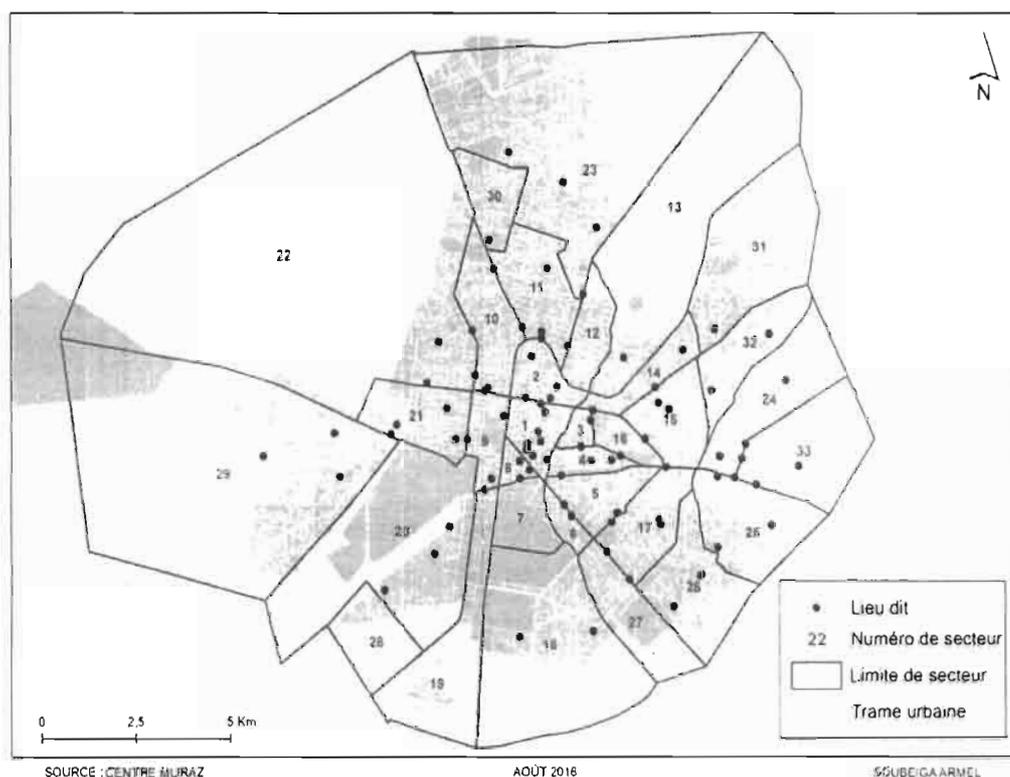
avec  $\pi_i, \pi_j$  et  $\Delta_{ij}$  strictement positifs.

## 2.3. Estimation d'un total selon le plan de sondage

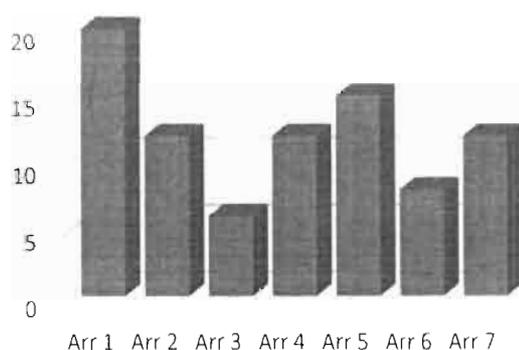
### 2.3.1. Base de sondage

Nous utilisons une base de sondage aréolaire de la commune urbaine de Bobo Dioulasso dont nous avons participé à l'élaboration durant le premier semestre de 2016. La base est composée de 85 lieux-dits. Les lieux-dits peuvent être des CSPS, des pharmacies, des mairies d'arrondissements, des antennes de réseau de téléphonie mobile, des bornes fontaines, etc. A chaque lieu-dit sont associés des caractéristiques telles que le nom, la nature, les coordonnées géographiques, le type de zone abritant le lieu-dit, la densité de la population 100m à la ronde autour du lieu-dit.

La figure ci-dessous présente la répartition spatiale des lieux-dits échantillonnés dans la commune urbaine de Bobo Dioulasso.



*Figure 1: Répartition spatiale des lieux-dits observés*



*Figure 2 : Répartition des lieux-dits par arrondissement*

### 2.3.1.1. La collecte de données

C'est un processus qui permet d'obtenir l'information nécessaire pour une étude. La pertinence et l'exactitude des données collectées définiront la qualité de l'étude. Elle se fait à travers un questionnaire préalablement établi par la structure chargée de l'enquête.

La collecte de données pour la présente recherche s'est déroulée en deux étapes :

- une première phase du 30 mai au 1<sup>er</sup> juin 2016,
- une seconde phase le 05 juillet 2016 à laquelle nous avons pris part.

Nous avons programmé le questionnaire sur support électronique ayant servi à cette collecte de données. A partir du logiciel CSPro 6.3 version ordinateur, nous avons créé un dictionnaire des

variables du questionnaire. De ces variables, nous isolons les fichiers permettant de réaliser la collecte sur support électronique : appareils dotés d'un système Android (ici tablettes).

➤ Le questionnaire

|   |
|---|
| <p><b>A. Identification</b><br/>Identifiant de l'enquêteur<br/>Identifiant du Lieu-Dit<br/>Date de collecte</p> <p><b>B. Localisation du Lieu-dit</b><br/>Dans quel arrondissement se trouve le lieu-dit ?<br/>Dans quel secteur se trouve le lieu-dit ?<br/>La Longitude du lieu-dit<br/>La Latitude du lieu-dit<br/>Le type de zone où se trouve le lieu-dit</p> <ol style="list-style-type: none"><li>1. Industrielle</li><li>2. Administrative</li><li>3. Commerciale</li><li>4. Habitation</li></ol> <p><b>C. Caractéristiques du Lieu-dit</b><br/>La nature du lieu-dit</p> <ol style="list-style-type: none"><li>1. Pharmacie</li><li>2. CSPS</li><li>3. Marie</li><li>4. Autre</li></ol> <p>Photographie (photo du lieudit)</p> |
|---|

### 2.3.1.2. Interpolation spatiale

L'interpolation spatiale est une procédure basée sur des algorithmes mathématiques, qui a permis d'estimer le nombre de ménages pour des sites non échantillonnés. Elle se déroule en 3 phases qui sont l'analyse exploratoire des données, l'analyse variographique et le krigeage. Ces méthodes ont été développées et décrites dans le rapport de stage d'Armel SOUBEIGA [4].

Cette modélisation avait pour objectif d'estimer le nombre de ménages  $z^*$  en une zone  $s$  non échantillonnée à partir du nombres de ménages  $z_i$ , observée auprès des lieux-dits échantillonnés.

Le nombre de ménage estimé est donné par la formule suivante :

$$z^* = \sum_{i=1}^n \lambda_i z_i$$

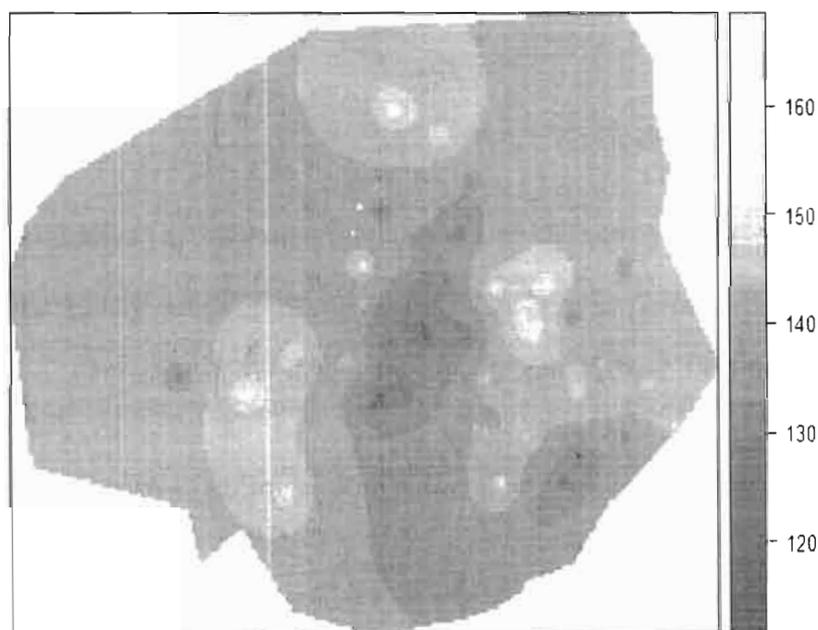
avec :

$\lambda_i$  : statistique de krigeage (interpolation spatiale)

$n$  : nombre de lieux-dits échantillonnés

Nous utiliserons  $z^*$  dans la suite du document pour le calcul des probabilités d'inclusion des ménages dans l'échantillon sélectionné.

La description de cette base a fait l'objet d'un autre rapport de stage.



*Figure 3 : Estimation des densités par interpolation spatiale*

Le nombre de ménages est estimé pour une aire d'un rayon de 100 m, soit d'une superficie de  $\pi$  hectares. On peut étendre la zone de couverture de l'UP en corrigeant le nombre d'habitants de la zone ainsi délimitée.

### 2.3.2. Plan de sondage

Nous nous plaçons dans le cas d'un échantillonnage à deux degrés. Nous décrivons de façon général le plan de sondage, une application à notre base aréolaire sera décrite par la suite.

On suppose que le plan de sondage vérifie les hypothèses classiques suivantes :

- l'invariance : à chaque fois qu'une unité primaire est sélectionnée dans l'échantillon du premier degré, le même plan est utilisé au second degré.
- l'indépendance : l'échantillonnage du second degré dans chaque UP se fait indépendamment du tirage de second degré dans les autres UP.

La population  $U_{GR}$  est constituée de  $M$  Unités Primaires (UP). Chaque UP  $u_h$  contient  $N_h$  unités secondaires (US). Soit  $m$  la taille de l'échantillon  $S_H$  d'UP sélectionnées selon un plan de sondage  $p_H$ . Si l'UP  $u_h$  est choisie dans  $S_H$ , on note  $S_h$  l'échantillon d'US sélectionné dans  $u_h$  selon un plan de sondage  $p_h$ . La taille de l'échantillon  $S_h$  est notée  $n_h$ .

Nous noterons  $U = \bigcup_{h=1}^M u_h$  l'ensemble des UP. L'échantillon final  $S$  d'US est donné par la réunion des  $S_h$ .

### 2.3.3. Probabilités d'inclusion

On note :

$\pi_{Hh}$  : la probabilité de sélectionner  $u_h$  dans l'échantillon  $S_H$  d'UP.

$\pi_{Hhk}$  : la probabilité de sélectionner conjointement  $u_h$  et  $u_k$  dans  $S_H$

$$\Delta_{Hhk} = \pi_{Hhk} - \pi_{Hh}\pi_{Hk}$$

$\pi_{i|h}$  : la probabilité de sélectionner l'US  $i$  appartenant à  $u_h$  dans  $S_h$

$\pi_{ij|h}$  : la probabilité de sélectionner les US  $i$  et  $j$  appartenant à  $u_h$  dans  $S_h$

$$\Delta_{ij|h} = \pi_{ij|h} - \pi_{i|h}\pi_{j|h}$$

En supposant un tirage aléatoire à probabilités égales et sans remise des unités primaires, on conclut que :

$$\pi_{Hh} = \frac{m}{M}$$

$$\pi_{Hhk} = \frac{m}{M} * \frac{m-1}{M-1}$$

De même, en supposant un sondage à probabilités égales et sans remise des unités secondaires à l'intérieur de leurs unités primaires, on conclut à :

$$\pi_{i|h} = \frac{n_h}{N_h}$$

$$\pi_{i|h} = \frac{n_h}{z_h^*}$$

$$\pi_{ij|h} = \frac{n_h}{N_h} * \frac{n_h-1}{N_h-1}$$

$$\pi_{ij|h} = \frac{n_h}{z_h^*} * \frac{n_h-1}{z_h^*-1}$$

Où  $z_h^*$  sera obtenu par estimation spatiale sur la base des données collectées par nos soins.

### 2.3.4. Estimateur de Horvitz-Thompson, précision et précision estimée

Le total  $T(U) = \sum_{i=1}^N Y_i$  de la variable  $Y$  sur  $U$  peut être estimé sans biais par son  $\pi$ -estimateur

$$\hat{T}_{HT}(S) = \sum_{i=1}^n \frac{y_i}{\pi_i} = \sum_{h=1}^m \sum_{i=1}^{n_h} \frac{y_i}{\pi_{Hh} \pi_{i|h}}$$

Pour l'estimation de la variance de  $\hat{T}(S)$ , les probabilités d'inclusions seront considérées comme non aléatoires.

Le total  $T(u_h)$  de la variable  $Y$  sur  $u_h$  est estimé sans biais par  $\hat{T}(S_h) = \sum_{i=1}^{n_h} \frac{y_i}{\pi_{i|h}}$ . On a donc la relation :

$$\hat{T}(S) = \sum_{h=1}^m \frac{\hat{T}(S_h)}{\pi_{Hh}}$$

La variance de  $\hat{T}(S)$  est estimée par

$$\begin{aligned} V(\hat{T}(S)) &= V\left(\sum_{h=1}^M \frac{T(u_h)}{\pi_{Hh}}\right) + \sum_{h=1}^M \frac{V(\hat{T}(S_h))}{\pi_{Hh}} \\ &= \underbrace{\sum_{h=1}^M \sum_{k=1}^M \frac{T(u_h) T(u_k)}{\pi_{Hh} \pi_{Hk}} \Delta_{Hhk}}_{V_{UP}} + \underbrace{\sum_{h=1}^M \frac{V(\hat{T}(S_h))}{\pi_{Hh}}}_{V_{US}} \end{aligned}$$

avec  $V(\hat{T}(S_h)) = \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} \frac{y_i}{\pi_{i|h}} \frac{y_j}{\pi_{j|h}} \Delta_{ij|h}$ . La variance se décompose en deux termes  $V_{UP}$  et  $V_{US}$ , respectivement associés aux premier et second degré de tirage. Si les probabilités d'inclusion  $\pi_{Hhk}$  et  $\pi_{ij|h}$  sont toutes strictement positives et non aléatoires, un estimateur sans biais de variance est donné par

$$\begin{aligned} \hat{V}(\hat{T}(S)) &= \hat{V}\left(\sum_{h=1}^m \frac{T(u_h)}{\pi_{Hh}}\right) + \sum_{h=1}^m \frac{\hat{V}(\hat{T}(S_h))}{\pi_{Hh}} \\ &= \underbrace{\sum_{h=1}^m \sum_{k=1}^m \frac{T(u_h) T(u_k)}{\pi_{Hh} \pi_{Hk}} \frac{\Delta_{Hhk}}{\pi_{Hhk}}}_{\hat{V}_A} + \underbrace{\sum_{h=1}^m \frac{V(\hat{T}(S_h))}{\pi_{Hh} * \pi_{ij|h}}}_{\hat{V}_B} \end{aligned}$$

avec  $\hat{V}(\hat{T}(S_h)) = \sum_{i=1}^{n_h} \sum_{j=1}^{n_h} \frac{y_i}{\pi_{i|h}} \frac{y_j}{\pi_{j|h}} \Delta_{ij|h}$

## 2.4. Autres méthodes d'estimation

Les formules précédentes sont applicables de façon générale. En pratique, il est souvent difficile de les utiliser vu leur complexité.

Compte-tenu de ces difficultés, il est fréquent d'utiliser des estimateurs simplifiés basés sur une approximation de la variance. Bien que (faiblement) biaisés, ces estimateurs sont généralement stables et présentent une erreur quadratique moyenne plus faible que celle de l'estimateur de Horvitz-Thompson.

### 2.4.1. Estimation par calage

Le calage est une méthode d'ajustement des poids d'un estimateur linéaire pondéré initial dans le but d'obtenir un nouvel estimateur linéaire qui estime parfaitement un certain nombre de totaux connus sur la population  $U$  (qu'on appellera totaux de calage ou totaux de contrôle). C'est une méthode qui permet d'intégrer de façon systématique l'information auxiliaire disponible.

Les poids obtenus par cette méthode sont appelés poids de calage.

La méthode du calage permet de construire une classe d'estimateurs pondérés linéaires, appelés estimateurs par calage.

Soit  $U = \{1, \dots, N\}$  une population de taille  $N$  à partir de laquelle un échantillon  $s$  de taille  $n$  est sélectionné selon un plan de sondage  $p(\cdot)$  dont les probabilités d'inclusion d'ordre un et deux sont données respectivement par  $\pi_k$  et  $\pi_{kl}$ . On s'intéresse à une variable d'intérêt  $y = (y_1, \dots, y_N)'$  en ayant pour objectif l'estimation de son total :

$$t_y = \sum_{k \in U} y_k$$

Un estimateur de  $t_y$  obtenu à partir de données complètes est le  $\pi$ -estimateur

$$\hat{t}_{y\pi} = \sum_{k \in S} d_k y_k$$

où  $d_k = 1/\pi_k$  est le poids d'échantillonnage de l'unité  $k$ .

Il est rare que l'on ne dispose pas d'une variable quantitative ou qualitative  $X$  dont la valeur/modalité est connue pour chacun des individus de la population (variable auxiliaire).

Lorsque nous disposons d'une information auxiliaire, nous devons chercher à l'utiliser dans le but d'obtenir des estimateurs plus précis que les estimateurs simples de la moyenne ou du total qui apparaissent dans le cadre du sondage. Cette information auxiliaire peut être utilisée au niveau de la construction de l'échantillon ou au niveau de l'expression de l'estimateur.

L'échantillon sélectionné sert normalement à effectuer une inférence sur la population. Pour passer de l'échantillon à la population, on utilise un poids attaché à chaque unité de l'échantillon : à chaque unité  $i$  de l'échantillon, on associe un poids  $w_i$ .

Le poids le plus simple permettant d'effectuer une estimation sans biais est le poids de sondage correspondant, dans le cas de l'estimation d'un total, à l'inverse de la probabilité d'inclusion  $p_i$  de l'unité  $i$  (estimateur de Horvitz-Thompson) : pour tout  $i \in U$ ,

$$w_i = \frac{1}{p_i}$$

Ces probabilités d'inclusion dépendant du plan de sondage utilisé, les poids de sondage dépendent eux aussi du plan de sondage mis en œuvre.

Le calage est une forme de redressement des estimations. Il ajuste les poids de sondage de sorte que les estimations soient calées sur des totaux (ou moyennes) connu(e)s : quel que soit l'échantillon sélectionné, on estime parfaitement ces totaux connus : on supprime l'erreur d'échantillonnage dans l'estimation des totaux connus.

#### **2.4.2. Estimation de la variance par bootstrap**

Le Bootstrap est sans doute la méthode d'estimation de précision par réplication la plus générale. Elle a été initialement proposée par Efron (1979) dans le cadre d'une population infinie.

Le principe des méthodes de rééchantillonnage est fondé sur l'utilisation répétée de l'échantillon de départ en le divisant, en supprimant, ou encore en répétant des observations pour obtenir plusieurs valeurs de l'estimateur permettant le calcul empirique d'une moyenne et d'une variance.

La méthode de rééchantillonnage, en anglais "bootstrap", consiste à tirer aléatoirement et avec remise, au sein de l'échantillon initial de taille  $n$ ,  $B$  "nouveaux" échantillons de même taille  $n$ . Cette opération est répétée un nombre  $B$  de fois, autant que nécessaire pour satisfaire un critère de convergence. L'estimateur est calculé sur chaque nouvel échantillon. Les variations entre les différentes valeurs de l'estimateur obtenues fournissent la variance.

## **3. Application**

### **3.1. Présentation de l'étude**

L'Equipe de recherche en Sociétés et Santé du Centre MURAZ a mené, en février 2011, une étude visant à identifier et à caractériser les incapacités chez les personnes âgées au Burkina Faso [5]. Il s'agissait de tirer un échantillon représentatif de ménages dans la ville, de recenser toutes les personnes âgées y vivant et ensuite de les soumettre à des outils d'évaluation des situations d'incapacités, le PRISMA7 et le SMAF. La sélection des ménages a été faite comme suit :

La population de personnes âgées de Bobo Dioulasso était estimée à 18130. De plus, on estimait à environ 0,056 la proportion de personnes en perte d'autonomie. Afin d'avoir un échantillon représentatif qui nous donnerait la prévalence des personnes âgées en perte d'autonomie avec une Erreur Relative de 0.3 ainsi qu'un effet de grappe de 1 avec un taux de non réponse de 10%, 727 personnes âgées avaient été prévus pour être sélectionnés dans la ville de Bobo Dioulasso. Les ménages étaient choisis dans les 25 secteurs de la ville(en 2012) proportionnellement à la taille en nombre de ménages.

Finalement, l'étude était réalisée en deux degrés. Le premier consistait à stratifier sur tous les 25 secteurs de la ville. Le second degré consistait ensuite à choisir les ménages à inclure avec un taux de sondage de 0,77%. Le choix des ménages était effectué selon la méthode des itinéraires

en respectant un pas systématique de 5 concessions, ce qui permet une approximation de l'aléatoire mais ne la garantit pas.

Il convient de noter que la variance issue d'une telle estimation sera très élevée étant donné les faibles probabilités d'inclusion qui ont été appliquées au second niveau.

Nous proposons une autre procédure de sondage qui aurait permis de réaliser la même étude avec des variances faibles pour les différents estimateurs.

## 3.2. Procédure de tirage des ménages

### ➤ PROBLEMATIQUE ET OBJECTIFS DE L'ETUDE

Le vieillissement de la population et l'accroissement du nombre de personnes âgées avec incapacités exercent une pression croissante sur la demande de services d'aide et de soins. Les dirigeants et les gestionnaires des centres de santé notamment le Centre MURAZ sont par conséquent amenés à revoir les mécanismes en place pour l'attribution des services et des prestations de soutien aux personnes en perte d'autonomie dans une perspective d'équité.

Objectif général de l'étude : identifier et à caractériser les incapacités chez les personnes âgées au Burkina Faso [6]

Objectifs spécifiques :

- i) déterminer le niveau d'autonomie fonctionnelle.
- ii) déterminer le degré de dépendance.

### ➤ POPULATION PARENTE

La population cible de l'étude est l'ensemble des personnes âgées vivant dans la commune urbaine de Bobo Dioulasso.

Les critères d'inclusion sont les suivants :

- i) être âgé de 60 ans ou plus (définition de l'OMS) [7];
- ii) résider dans la commune urbaine de Bobo Dioulasso.

### ➤ PROPOSITION DE QUESTIONNAIRE

Plusieurs outils d'évaluation de l'autonomie fonctionnelle existent. Dans le cadre de cette étude nous proposons de n'utiliser que le Prisma-7 et la grille d'évaluation SMAF :

- [8]le Prisma-7 (Programme de Recherche pour l'Intégration des Services pour le Maintien de l'Autonomie) est un outil de repérage permettant d'identifier les personnes en perte d'autonomie. Il s'agit d'un questionnaire de sept questions auxquelles il faut répondre par oui ou par non.

1. Avez-vous plus de 85 ans ?
2. Sexe masculin ?
3. En général, est-ce que des problèmes de santé vous obligent à limiter vos activités ?
4. En général, est-ce que des problèmes de santé vous obligent à rester à la maison ?
5. Avez-vous besoin de quelqu'un pour vous aider régulièrement ?
6. Pouvez-vous compter sur une personne qui vous est proche en cas de besoin ?
7. Utilisez-vous régulièrement une canne ou une marchette ou un
8. fauteuil roulant pour vous déplacer ?

*Tableau 1 : Questionnaire PRISMA-7*

- la grille d'évaluation SMAF évalue 29 fonctions couvrant les activités de la vie quotidienne, la mobilité, la communication, les fonctions mentales, les tâches domestiques. Les fonctions sont cotées sur une échelle à cinq degrés : 0 (autonome) ; -1 (surveillance ou stimulation requises) ; -2 (aide partielle requise) ; -3 (aide complète requise) ; une cote intermédiaire de -0,5 ou de -1,5 est utilisée pour certains items lorsqu'une activité est réalisée de façon autonome mais avec difficulté.

|  |   |
|--|---|
| <b>Activités de la vie quotidienne</b> | <b>Se nourrir</b>                             |
|  | S'habiller                                    |
|  | Se laver                                      |
|  | Entretenir sa personne                        |
|  | Fonction vésicale                             |
|  | Fonction anale                                |
|  | Utiliser toilette                             |
| <b>Mobilité</b>                        | Transferts                                    |
|  | Marcher à l'intérieur                         |
|  | Installer prothèse ou orthèse                 |
|  | Se déplacer en fauteuil roulant à l'intérieur |
|  | Utiliser les escaliers                        |
| <b>Communication</b>                   | Circuler à l'extérieur                        |
|  | Voir  |
|  | Entendre                                      |
| <b>Fonctions cognitives</b>            | Parler  |
|  | Jugement                                      |
|  | Comportement                                  |
|  | Orientation                                   |
|  | Mémoire                                       |
| <b>Activités de la vie domestique</b>  | Compréhension                                 |
|  | Utiliser le téléphone maison                  |
|  | Entretenir la maison                          |
|  | Faire la lessive                              |
|  | Préparer les repas                            |
|  | Faire les courses                             |
|  | Utiliser les moyens de transport              |
|  | Prendre ses médicaments                       |

Tableau 2: Variables de la grille du SMAF

En plus de ces questions d'ordre évaluative de l'autonomie, d'autres questions d'intérêts pourraient être ajoutées.

➤ ECHANTILLONNAGE

La procédure de sondage utilisée est celle décrite dans la partie Matériels et méthodes. Il s'agit d'un échantillonnage à deux degrés s'appuyant sur une base aréolaire. Les UP et les US possèdent la même probabilité d'inclusion dans l'échantillon.

Un échantillon de 727 personnes âgées est sélectionné pour l'étude comme dans l'étude réalisé en Février 2011. A raison d'une personne âgée par ménage, le tirage se limitera donc à 727 ménages.

➤ TIRAGE DES ECHANTILLONS

- Premier degré

Le territoire de la commune urbaine de Bobo-Dioulasso est représenté par des disques de  $\pi$  hectares chacun. Le tirage est aléatoire sans remise, à probabilités égales. Il s'effectue à l'aide des coordonnées géographiques des lieux-dits. L'échantillon est réparti entre les différents lieux-dits proportionnellement à leur taille en nombre de ménages. Le nombre de lieux-dits tirés est de  $m$  (nombre fixé arbitrairement).

- Second degré

Les ménages sont sélectionnés selon la méthode des itinéraires avec un pas arbitraire pour ménages. Tout ménage dans lequel aucune personne âgée n'est trouvée sera automatiquement remplacé par le ménage le plus proche dans lequel peuvent être trouvées des personnes âgées.

# 4. Discussion

## 4.1. Avantages et limites

La procédure ainsi développée est peu onéreuse comparativement à un recensement. Le coût total de ce sondage est inférieur à celui d'un recensement, ce dernier exigeant des ressources humaines, financières, logistiques et matérielles considérables. Un échantillon probabiliste, s'il est sélectionné et analysé comme il convient, permettra d'obtenir des résultats exacts et fiables qui pourront servir de base à des extrapolations à l'ensemble de la population.

Elle fournira des estimateurs efficaces comparativement à l'étude réalisée par le centre MURAZ en février 2011. Un estimateur efficace est un estimateur dont la variance est le plus faible possible.

Une des limites de cette étude est qu'en plus du sondage par degré, on aurait pu utiliser un autre critère de stratification. Aussi, au lieu des probabilités égales, des probabilités inégales pourraient être mises en œuvre. Cela aurait permis de tendre vers encore plus d'efficacité mais les formules de calcul ne seraient plus les mêmes.

Les formules que nous proposons n'ont pas été implémentées. C'aurait été meilleur si nous avions pu proposer un programme sous le logiciel R qui calculerait les différentes statistiques définis dans l'ouvrage : les estimateurs et leurs variances associés. Nous aurions également souhaité que l'étude soit réalisée.

## 4.2. Perspectives

[9]Le présent travail de recherche a été accepté et présenté lors de la conférence internationale de Statistique Appliquée pour le Développement en Afrique (SADA 2016) qui s'est déroulé du 28 novembre au 3 décembre 2016 à Cotonou au Bénin. Cette conférence est un lieu privilégié de discussions et d'échanges sur les différents aspects de la statistique et de la modélisation stochastique, tant théoriques que pratiques, en relation avec les problèmes du développement en Afrique. Elle mise sur une triple synergie : rencontre entre professionnels et chercheurs issus de divers milieux (université, organisme de recherche scientifique, entreprise, organismes gouvernementaux et intergouvernementaux, etc.) et disciplines, rencontre entre les différentes cultures scientifiques (francophone, anglophone, lusophone, hispanophone, etc.), qui s'expriment sur le continent africain, rencontre enfin entre statisticiens travaillant en Afrique et à l'extérieur de l'Afrique.

Dans le cadre de la collecte électronique de données, nous avons contribué à la mise à jour des données de la plateforme OpenStreetMap (OSM). OSM est un projet international fondé en 2004 dans le but de créer une carte libre du monde. Il collecte des données dans le monde entier sur les routes, voies ferrées, les rivières, les forêts, les bâtiments, etc. Grâce aux données recueillis lors de nos sorties terrains, nous avons contribué à répertorié quelques sites géographiques de la commune urbaine de Bobo et mettre à jour des données obsolètes sur la carte OpenStreetMap en ligne.

A long terme, le système de sondage établi pourrait être adapté et utiliser pour estimer la population de la commune urbaine de Bobo Dioulasso de façon ponctuelle.

## 5. Conclusion

Ce stage a été très enrichissant pour nous car il nous a permis d'une part de découvrir le secteur de la recherche fondamentale et appliquée dans le domaine des sciences de la santé, ses acteurs ainsi que ses contraintes. Nous avons participé concrètement à ses enjeux au travers de nos missions variées comme celle de la collecte de données, l'analyse de données, l'interpolation spatiale et l'élaboration de plan de sondage.

D'autre part, nous avons découvert notre appétence pour les missions créatives entre autres, l'utilisation de méthodes statistiques et informatique en vue de concevoir ou de résoudre de façon effective des problèmes concrets. Nous préférons donc nous orienter dans un futur proche vers les métiers de l'ingénierie statistique.

L'une des ambitions du Centre MURAZ est de devenir un pôle régional d'excellence et d'innovation dans la recherche au service de la santé publique. Il a donc besoin de nouvelles méthodologies de sondage à la fois probabilistes et économiques pour demeurer compétitif. Nous sommes heureux d'avoir pu apporter notre pierre à cet édifice de méthodologies au travers du développement et de la validation d'un système de sondage à base aréolaire dans la commune urbaine de Bobo Dioulasso.

## Bibliographie

- [1] J. Vial, «Définition de la validation de méthode et outils associés».
- [2] INSD, «Monographie de la Commune Urbaine de Bobo Dioulasso,» Bobo Dioulasso, Décembre 2009.
- [3] Institut National de la Statistique et de la Démographie, «Recensement Général de la Population et de l'Habitation (RGPH),» 2006.
- [4] A. Soubeiga, «Statistique Spatiale : Cartographie et modélisation spatiale par krigeage,» 2016.
- [5] A. Berthe, Le système burkinabè de maintien des personnes âgées en autonomie fonctionnelle à domicile analyse centrée sur les acteurs de la ville de Bobo-Dioulasso., Louvain: UCL, 2013.
- [6] A. BERTHÉ, L. BERTHÉ-SANOU, S. SOMDA et e. al., «The key actors maintaining elders in functional autonomy in Bobo-Dioulasso (Burkina Faso).,» *BMC public health*, vol. 14, n° 11, p. 1, 2014.
- [7] A. Dao, «Les Caractéristiques Sociodémographiques des personnes âgées dans la ville de Bobo Dioulasso,» 2014.
- [8] Centre d'expertise en santé de Sherbrooke, «Guide d'utilisation du questionnaire PRISMA-7,» 2010.
- [9] S. Somda, D. E. Sanou et A. Soubeiga, «Defining a new sampling system in African urban statement based on spatial estimation,» chez SADA, Cotonou, 2016.
- [10] L. Rouvière, «Enquête et Sondage,» 2008-2009.
- [11] G. Chauvet, «Techniques avancées d'échantillonnage, Estimation de variance,» 28 janvier 2014.
- [12] J.-J. DROESBEKE, B. FICHET et P. TASSI, Les Sondages, 1987.
- [13] M. d. l'agriculture, «Méthodologie de l'enquête nationale sur les statistiques agricoles,» 1993.

## Annexe

Tableau 3 : Lieux-dits échantillonnés

| Num. | Arr. | Secteur | Longitude | Latitude | Zone           | Nature    | Nombre ménages |
|------|------|---------|-----------|----------|----------------|-----------|----------------|
| 1    | 5    | 17      | -4.26936  | 11.15995 | Habitation     | CSPS      | 22             |
| 2    | 5    | 17      | -4.26902  | 11.15893 | Habitation     | Pharmacie | 44             |
| 3    | 5    | 17      | -4.26761  | 11.17169 | Habitation     | Pharmacie | 25             |
| 4    | 5    | 5       | -4.27942  | 11.1617  | Habitation     | Mairie    | 12             |
| 5    | 5    | 17      | -4.28078  | 11.15954 | Habitation     | Pharmacie | 46             |
| 6    | 5    | 5       | -4.28213  | 11.1527  | Habitation     | Pharmacie | 28             |
| 7    | 7    | 21      | -4.31899  | 11.17741 | Habitation     | CSPS      | 44             |
| 8    | 7    | 21      | -4.31623  | 11.17725 | Habitation     | Pharmacie | 38             |
| 9    | 6    | 9       | -4.30729  | 11.18251 | Habitation     | Pharmacie | 13             |
| 10   | 1    | 1       | -4.30166  | 11.17532 | Commerciale    | Pharmacie | 0              |
| 11   | 2    | 10      | -4.30148  | 11.17646 | Commerciale    | Pharmacie | 0              |
| 12   | 1    | 1       | -4.29811  | 11.17717 | Commerciale    | Pharmacie | 0              |
| 13   | 6    | 9       | -4.31194  | 11.16618 | Habitation     | Pharmacie | 7              |
| 14   | 5    | 6       | -4.2924   | 11.16334 | Habitation     | Pharmacie | 26             |
| 15   | 1    | 8       | -4.30307  | 11.16899 | Habitation     | Pharmacie | 8              |
| 16   | 1    | 4       | -4.29303  | 11.1697  | Administrative | Mairie    | 0              |
| 17   | 1    | 4       | -4.28585  | 11.17292 | Habitation     | Pharmacie | 17             |
| 18   | 1    | 1       | -4.28088  | 11.17319 | Habitation     | CSPS      | 46             |
| 19   | 4    | 16      | -4.27865  | 11.17411 | Habitation     | Pharmacie | 33             |
| 20   | 1    | 4       | -4.28832  | 11.17609 | Habitation     | Pharmacie | 26             |
| 21   | 1    | 3       | -4.28541  | 11.18398 | Commerciale    | Pharmacie | 9              |
| 22   | 4    | 15      | -4.27265  | 11.17782 | Habitation     | Pharmacie | 29             |
| 23   | 5    | 25      | -4.25092  | 11.16937 | Habitation     | Pharmacie | 29             |
| 24   | 4    | 24      | -4.24925  | 11.17359 | Habitation     | CSPS      | 16             |
| 25   | 4    | 24      | -4.24836  | 11.17671 | Habitation     | Pharmacie | 36             |
| 26   | 4    | 24      | -4.25458  | 11.17409 | Habitation     | Mairie    | 38             |
| 27   | 5    | 6       | -4.29057  | 11.16092 | Habitation     | CSPS      | 22             |
| 28   | 6    | 20      | -4.32045  | 11.15822 | Habitation     | Pharmacie | 34             |
| 29   | 6    | 20      | -4.32408  | 11.15201 | Administrative | CSPS      | 0              |
| 30   | 7    | 29      | -4.33476  | 11.17839 | Habitation     | Pharmacie | 45             |
| 31   | 7    | 29      | -4.36553  | 11.17351 | Habitation     | CSPS      | 7              |
| 32   | 7    | 21      | -4.32115  | 11.1841  | Habitation     | Pharmacie | 77             |
| 33   | 7    | 22      | -4.31427  | 11.1913  | Habitation     | Pharmacie | 48             |
| 34   | 7    | 29      | -4.31484  | 11.20141 | Habitation     | Pharmacie | 55             |
| 35   | 7    | 22      | -4.32319  | 11.19886 | Habitation     | Pharmacie | 24             |
| 36   | 7    | 22      | -4.32585  | 11.18977 | Habitation     | Pharmacie | 22             |
| 37   | 7    | 22      | -4.33311  | 11.18059 | Habitation     | Mairie    | 37             |
| 38   | 2    | 10      | -4.31192  | 11.18794 | Habitation     | Pharmacie | 26             |
| 39   | 2    | 10      | -4.31118  | 11.18852 | Habitation     | CSPS      | 24             |
| 40   | 1    | 2       | -4.30198  | 11.1865  | Commerciale    | Pharmacie | 16             |
| 41   | 1    | 2       | -4.30059  | 11.19575 | Habitation     | Pharmacie | 17             |
| 42   | 2    | 11      | -4.30284  | 11.20207 | Habitation     | Pharmacie | 36             |

|    |   |    |          |          |                |           |    |
|----|---|----|----------|----------|----------------|-----------|----|
| 43 | 2 | 11 | -4.29676 | 11.21503 | Habitation     | CSPS      | 30 |
| 44 | 2 | 10 | -4.3098  | 11.21494 | Habitation     | Pharmacie | 12 |
| 45 | 3 | 12 | -4.28779 | 11.20962 | Habitation     | CSPS      | 4  |
| 46 | 3 | 12 | -4.28474 | 11.22399 | Habitation     | Pharmacie | 16 |
| 47 | 2 | 10 | -4.29793 | 11.19991 | Habitation     | Mairie    | 15 |
| 48 | 1 | 2  | -4.29569 | 11.18658 | Habitation     | Pharmacie | 30 |
| 49 | 1 | 1  | -4.29711 | 11.18357 | Habitation     | CSPS      | 7  |
| 50 | 1 | 2  | -4.29807 | 11.1853  | Commerciale    | Pharmacie | 11 |
| 51 | 6 | 9  | -4.3101  | 11.16887 | Habitation     | Mairie    | 3  |
| 52 | 1 | 8  | -4.30316 | 11.17282 | Habitation     | Pharmacie | 20 |
| 53 | 1 | 8  | -4.30083 | 11.17089 | Administrative | CSPS      | 0  |
| 54 | 1 | 2  | -4.29431 | 11.18901 | Habitation     | CSPS      | 32 |
| 55 | 1 | 3  | -4.28588 | 11.18185 | Habitation     | CSPS      | 16 |
| 56 | 3 | 14 | -4.2702  | 11.18906 | Habitation     | Pharmacie | 53 |
| 57 | 4 | 15 | -4.26702 | 11.18409 | Habitation     | Pharmacie | 49 |
| 58 | 4 | 15 | -4.26955 | 11.18546 | Habitation     | CSPS      | 38 |
| 59 | 4 | 24 | -4.24573 | 11.16781 | Habitation     | Pharmacie | 34 |
| 60 | 5 | 25 | -4.2551  | 11.16953 | Habitation     | CSPS      | 40 |
| 61 | 3 | 13 | -4.27795 | 11.19577 | Habitation     | Mairie    | 55 |
| 62 | 2 | 11 | -4.29173 | 11.19811 | Commerciale    | Pharmacie | 12 |
| 63 | 1 | 1  | -4.29882 | 11.17922 | Commerciale    | Pharmacie | 0  |
| 64 | 1 | 1  | -4.29996 | 11.174   | Commerciale    | Pharmacie | 0  |
| 65 | 1 | 1  | -4.2968  | 11.17304 | Commerciale    | Pharmacie | 0  |
| 66 | 2 | 11 | -4.29791 | 11.20128 | Habitation     | Pharmacie | 33 |
| 67 | 5 | 17 | -4.27642 | 11.14682 | Habitation     | Pharmacie | 55 |
| 68 | 6 | 28 | -4.33598 | 11.14419 | Habitation     | Autre     | 54 |
| 69 | 6 | 18 | -4.28516 | 11.13572 | Habitation     | Autre     | 22 |
| 70 | 6 | 18 | -4.30333 | 11.13417 | Habitation     | Autre     | 18 |
| 71 | 5 | 27 | -4.26578 | 11.14088 | Habitation     | Autre     | 8  |
| 72 | 5 | 25 | -4.25491 | 11.15397 | Habitation     | Autre     | 1  |
| 73 | 5 | 26 | -4.25908 | 11.14775 | Habitation     | Autre     | 8  |
| 74 | 5 | 25 | -4.24204 | 11.159   | Habitation     | Autre     | 13 |
| 75 | 4 | 33 | -4.23548 | 11.17201 | Habitation     | Autre     | 38 |
| 76 | 4 | 24 | -4.23865 | 11.19076 | Habitation     | Autre     | 32 |
| 77 | 4 | 32 | -4.25657 | 11.18837 | Habitation     | Autre     | 4  |
| 78 | 3 | 31 | -4.26372 | 11.19732 | Habitation     | Autre     | 53 |
| 79 | 3 | 31 | -4.25586 | 11.20194 | Habitation     | Autre     | 0  |
| 80 | 4 | 32 | -4.24265 | 11.20106 | Habitation     | Autre     | 17 |
| 81 | 2 | 30 | -4.31087 | 11.22103 | Habitation     | Autre     | 33 |
| 82 | 2 | 23 | -4.3062  | 11.24045 | Habitation     | Autre     | 57 |
| 83 | 2 | 23 | -4.2929  | 11.23394 | Habitation     | Autre     | 49 |
| 84 | 7 | 29 | -4.34827 | 11.17864 | Habitation     | Autre     | 33 |
| 85 | 7 | 29 | -4.34711 | 11.16894 | Habitation     | Autre     | 57 |