

Tiral Sidi

# **Mon plan de sondage en 9 questions**

**Publibook**

Retrouvez notre catalogue sur le site des Éditions Publibook :

<http://www.publibook.com>

Ce texte publié par les Éditions Publibook est protégé par les lois et traités internationaux relatifs aux droits d'auteur. Son impression sur papier est strictement réservée à l'acquéreur et limitée à son usage personnel. Toute autre reproduction ou copie, par quelque procédé que ce soit, constituerait une contrefaçon et serait passible des sanctions prévues par les textes susvisés et notamment le Code français de la propriété intellectuelle et les conventions internationales en vigueur sur la protection des droits d'auteur.

Éditions Publibook  
14, rue des Volontaires  
75015 PARIS – France  
Tél. : +33 (0)1 53 69 65 55

IDDN.FR.010.0114061.000.R.P.2009.030.40000

Cet ouvrage a fait l'objet d'une première publication aux Éditions Publibook en 2010

A mes enfants Raina, Zeynab et Fayçal



## Remerciements

Mes remerciements vont à mes professeurs Yves Aragon et Anne Ruiz-Gazen qui ont suscité et développé en moi l'intérêt aux techniques de sondage. J'adresse également mes remerciements à Ghislaine Delaine, statisticienne principale à la Banque mondiale et responsable des enquêtes sur la pauvreté dans plusieurs pays africains et à Alphonse Yameogo, statisticien expert dans la conception et le traitement des enquêtes sur la pauvreté pour les précieux conseils et les suggestions qu'ils m'ont prodigués tout au long de la rédaction du présent livre.

## Avant-propos

Mon expérience dans l'élaboration des enquêtes et celle acquise à travers l'appui que j'ai apporté ces dernières années à certains instituts de statistique en Afrique, m'ont amené à relever des insuffisances dans la planification de l'élaboration des plans de sondage de ces enquêtes, ce qui est à l'origine de la rédaction de ce livre.

Ce livre propose une démarche simple et pratique pour tous ceux qui veulent élaborer un plan de sondage pour mener des enquêtes. Les différentes étapes du plan de sondage sont passées en revue de façon chronologique sous forme de questions et étayées par des exemples fictifs ou quand cela est possible, puisés dans des enquêtes réelles élaborées dans certains pays africains (Burkina Faso, Niger, Cap vert, Mauritanie, République Centrafricaine, etc.). La réponse à chaque question posée permet de résoudre une étape du plan de sondage et l'ensemble des réponses à ces questions constitue le plan de sondage pour l'enquête en question.

La simplicité pratique du livre, volontairement voulu par l'auteur n'exclut pas quelques démonstrations mathématiques minimales pour justifier les fondements théoriques de certains principaux résultats. Le lecteur pourra, s'il le souhaite, ignorer ces démonstrations pour s'intéresser uniquement aux résultats les plus importants et aux applications qui sont parfois illustrés par des exemples. En guise de référence rapide, le lecteur peut consulter des encadrés résumant l'essentiel au début de chaque question, sauf pour les questions 5 et 8 pour lesquelles ces résumés sont à la fin sous forme de tableaux.

Le lecteur intéressé par la théorie des sondages pourra consulter l'importante littérature en la matière comportant d'excellents livres dont certains font le tour des pratiques de sondages et d'autres sont beaucoup plus spécialisés sur certains aspects particuliers (voir bibliographie).

Il est important de souligner que l'élaboration du plan de sondage d'une enquête donnée est une étape primordiale qu'il faut conduire avec beaucoup de minutie. Le meilleur des plans de sondage ne garantit cependant pas à lui tout seul des résultats fiables pour une enquête. Il constitue seulement l'une des faces d'une médaille dont l'autre face est l'élaboration du questionnaire, la collecte et le traitement. En effet, les résultats d'une enquête peuvent être entachés par des erreurs d'observation (appelées aussi erreurs de réalisation) qui peuvent provenir d'une mauvaise organisation de l'enquête et d'un contrôle et d'une supervision insuffisants.

Tiral Sidi

# Introduction

Dans sa dimension la plus élémentaire, le sondage vise à estimer un total, une moyenne d'une variable d'étude sur une population. Citons par exemple, le nombre de chômeurs à *Bobo-Dioulasso*<sup>1</sup>, la dépense moyenne de thé et de sucre des ménages dans la province de l'*Oudalan*<sup>2</sup> ou la quantité moyenne de dolo consommée par personne dans le *Goulmou*<sup>3</sup> ou dans le *Bobodougou*<sup>4</sup>).

La théorie des sondages pose globalement deux types de problèmes qui sont liés. Il s'agit de la théorie de l'échantillonnage et de la théorie de l'estimation et des tests. En effet, la méthode d'échantillonnage utilisée a une influence sur les estimations obtenues.

C'est pourquoi, il est très important de ne pas perdre de vue cette liaison lors de l'analyse et de la publication des résultats d'une enquête donnée issue d'un plan de sondage probabiliste. Certaines conclusions faites et non validées par un test statistique sont hasardeuses et ne prennent pas en compte le fondement théorique même du plan de sondage probabiliste.

Nous nous intéressons dans ce livre aux différentes étapes de l'élaboration d'un plan de sondage des enquêtes, particulièrement des grandes enquêtes, dans l'espoir de mettre entre les mains des concepteurs de ces enquêtes un

---

<sup>1</sup> Ville du Burkina Faso

<sup>2</sup> Province du Burkina Faso

<sup>3</sup> Région des gourmantché (Parents à plaisanterie)

<sup>4</sup> Région des bobos (Autre parent à plaisanterie par assimilation)



document pratique, simple et quasi complet en termes d'étapes pour la conception et la planification d'un plan de sondage probabiliste.

Le constat fait dans certains pays montre que l'élaboration des plans de sondage des enquêtes menées par certaines structures de statistique est conduite par des statisticiens qui ne sont pas toujours spécialisés dans le domaine du sondage. Ils sont surtout des statisticiens d'enquêtes spécialisés principalement à deux niveaux : la conception des questionnaires, l'organisation de la collecte. Ils ont, certes une formation générale en la matière et de fait l'élaboration du plan de sondage n'est pas cernée dans toutes ses dimensions. Le constat est qu'un plan de sondage d'une enquête non élaboré par un appui extérieur, est parfois reproduit pour d'autres enquêtes sans tenir compte systématiquement des objectifs thématiques, des domaines d'étude ou de l'existence de variables auxiliaires à même d'améliorer la précision de l'échantillon.

D'autre part, la taille de l'échantillon déterminée pour certaines enquêtes ne tient pas toujours compte des effets de structure ou effet de sondage (*design effect en anglais*) et des besoins de comparaison dans le temps des indicateurs que l'enquête cherche à estimer. La mesure de l'impact de politiques de développement dans une zone donnée demande que l'on tienne compte aussi bien de la signification statistique que de la puissance statistique, ce qui n'est en général pas le cas dans la plupart des plans de sondage de ces enquêtes. Dans une telle situation l'analyse de certains résultats, conduit parfois à des polémiques entre les utilisateurs et que les statisticiens ont du mal à soutenir en l'absence d'une validation statistique robuste c'est-à-dire prenant en compte la puissance statistique.

Enfin, pour élaborer un plan de sondage, le statisticien est amené souvent à consulter plusieurs livres qui ne sont pas toujours disponibles, pour faire un tour d'horizon de ce qui

lui est nécessaire. Aussi, si l'essentiel de ce qui est utile pour faire un plan de sondage est dans un seul document, le travail devient pratiquement plus facile.

L'élaboration de plans de sondage adaptés à l'organisation et à l'accessibilité de la population par l'enquête est une tâche importante et délicate qui relève de spécialistes.

En sondage, il existe plusieurs méthodes de choisir les individus à observer. En effet, en plus des sondages probabilistes, il existe des méthodes non probabilistes ou méthodes par choix raisonnés (méthodes des unités-types, méthode des itinéraires, échantillonnage par quotas, etc.). Dans ce livre, nous nous intéresserons principalement aux enquêtes par sondage probabiliste.

Il ne s'agit pas ici de faire un exposé détaillé sur les sondages probabilistes. Le lecteur intéressé pourra se référer à d'excellents livres (voir bibliographie) qui existent. Nous proposons simplement l'élaboration d'un plan de sondage en répondant à neuf questions pour permettre une planification chronologique des étapes les plus importantes. Le plan de sondage de l'enquête concernée sera entièrement élaboré une fois que des réponses seront apportées à toutes ces questions. Il est vivement conseillé de capitaliser à chaque fois ces réponses et leur justification technique dans un document qui peut devenir une référence méthodologique pour toutes les enquêtes similaires et comme guide pratique pour l'élaboration du plan de sondage de n'importe autre enquête dans le pays.

Après avoir répondu à la « *Question 0 : Pourquoi un plan de sondage ?* », c'est-à-dire, après que le sondage ait été choisi comme méthode de collecte, alors quelles sont les différentes questions auxquelles doit répondre le concepteur du plan de sondage ? Les neuf questions proposées constituent les différentes étapes successives de la

conception d'un plan de sondage. Elles sont énumérées ci-après :

- Question 1.        Quels sont les objectifs thématiques de l'enquête ?*
- Question 2.        Quels sont les domaines d'étude de l'enquête ?*
- Question 3.        De quel budget dispose l'enquête ?*
- Question 4.        Existe-il une base de sondage ?*
- Question 5.        Quelle taille pour l'échantillon ?*
- Question 6.        Comment tirer les unités de l'échantillon ?*
- Question 7.        Comment améliorer la précision des estimations ?*
- Question 8.        Quels sont les estimateurs ?*
- Question 9.        Précision et validation des résultats : que faut-il faire ?*

Les réponses à ces questions feront l'objet des chapitres suivants. Elles seront donc toutes traitées et certaines seront illustrées par des exemples fictifs ou tirés d'enquêtes réelles de certains pays africains.

Nous terminerons en faisant quelques recommandations sur la conception du questionnaire et le contrôle de la collecte dont l'importance sur la qualité des données n'est plus à démontrer.

## Table des matières

Question 0 : Pourquoi un sondage ? .....	20
Question 1 : Quels sont les objectifs thématiques de l'enquête ? .....	24
Question 2 : Quels sont les domaines d'étude de l'enquête ? .....	30
Question 3 : De quel budget dispose l'enquête ? .....	36
Question 4 : Existe-t-il une base de sondage ? .....	40
4.1. Unité statistique .....	45
4.2. Unités primaires éligibles pour le tirage .....	47
Question 5 : Quelle taille pour l'échantillon ? .....	55
5.1 Rappels .....	56
5.1.1 Coefficient de variation .....	56
5.1.2 Inégalité de Markov .....	57
5.1.3 Inégalité de Bienaymé-Tchebychev .....	57
5.1.4 Loi des grands nombres .....	58
5.1.5 Théorème central-limite .....	60
5.1.6 Intervalle de confiance .....	61
5.1.7 Marge d'erreur absolue .....	61
5.1.8 Marge d'erreur relative .....	62
5.2 Eléments qui influencent la taille de l'échantillon .....	64
5.2.1 Effet de sondage .....	64
5.2.2 Non-réponses .....	66
5.2.3 Signification statistique .....	67
5.2.4 Puissance statistique .....	68
5.3 Taille de l'échantillon pour une enquête ponctuelle .....	71
5.3.1 Contrôle de la marge d'erreur absolue .....	71
5.3.2 Contrôle de la marge d'erreur relative .....	74
5.3.3 Contrôle du coefficient de variation de la variable d'intérêt .....	81
5.3.4 Taille de l'échantillon des unités primaires .....	86

5.4 Taille de l'échantillon pour une enquête d'impact.....	87
5.4.1 <i>Proportion</i> .....	87
5.4.2 <i>Moyenne</i> .....	88
5.5 Effectif des groupes de traitement en biostatistique.	89
5.6 Exemples de calcul de taille d'échantillon.....	93
5.6.1 <i>Contrôle de la marge d'erreur absolue pour une moyenne</i>	93
5.6.2 <i>Contrôle de la marge d'erreur relative pour une moyenne</i>	94
5.6.3 <i>Contrôle de la marge d'erreur absolue pour une proportion</i>	95
5.6.4 <i>Contrôle de la marge d'erreur relative pour une proportion</i>	96
5.6.5 <i>Contrôle du coefficient de variation</i> .....	97
5.7 Résumé de la détermination de la taille de l'échantillon	98
Question 6 : Comment tirer les unités de l'échantillon ?..	102
6.1 Tirage aléatoire simple .....	102
6.2 Tirage aléatoire systématique .....	107
Question 7 : Comment améliorer la précision des estimations?	115
7.1. Utilisation de variables auxiliaires .....	115
7.2. Stratification .....	116
7.2.1 <i>Objectif de précision locale</i> .....	119
7.2.2 <i>Objectif de précision globale</i> .....	120
7.2.3 <i>Répartition de la taille de l'échantillon entre les strates</i>	120
7.2.4 <i>Estimateurs</i> .....	124
7.3. Traitement des non-réponses .....	130
Question 8 : Quels sont les estimateurs ? .....	135
8.1. Probabilité d'inclusion .....	137
8.2. Plans de sondage élémentaires.....	138
8.2.1 Tirage sans remise.....	139
8.2.4 <i>Proportion</i> .....	158
8.2.5 <i>Ratio</i> .....	161
8.3 Tableau 8.2 : récapitulatif des principaux estimateurs	164
8.3.1 <i>Total et moyenne</i> .....	164
8.3.2 <i>Proportion et ratio</i> .....	165
8.4 Estimation dans un domaine .....	166

8.4.1 <i>Domaine d'étude a priori</i> .....	166
8.4.2 <i>Domaine d'étude a posteriori</i> .....	166
8.5 : Plans de sondage à plusieurs degrés .....	169
8.5.1 <i>Sondage en grappes</i> .....	169
8.5.2 <i>Cas des sondages à deux degrés</i> .....	170
8.5.3 <i>Espérances conditionnelles</i> .....	172
8.5.4 : <i>Calcul des poids de sondage</i> .....	178
Question 9 : Précision et validation des résultats : que doit-on faire ? .....	182
9.1 : Erreurs d'échantillonnages.....	183
9.2 Comparaison de résultats .....	190
Qualité des données : questionnaire et contrôle.....	195
Evitons un simple «réchauffement» des anciens questionnaires.....	195
Systématisation et exhaustivité du contrôle .....	198
Annexes .....	202
Quelques définitions.....	204
Concepts et définitions .....	207
Erreurs non dues à l'échantillonnage.....	210
Exemples de méthodes non probabilistes.....	212
Bibliographie .....	214