



**ÉTUDE COMBINATOIRE, PROBABILISTE ET COMPARATIVE DES METHODES
D'ECHANTILLONNAGES APPLIQUÉES AUX STRUCTURES SPATIALES URBAINES : CAS
DE L'HABITAT ET DE LA MOBILITÉ URBAINE**

Par

WANE MUNDONI BONGUON

Professeur Associé à l'Institut Supérieur d'Architecture et d'Urbanisme, RDC

Résumé :

Le présent travail s'inscrit dans le cadre d'une approche expérimentale, en vue de caractériser deux méthodes d'échantillonnage. Pour ce faire, des données de problématiques urbaines ont fait l'objet de comparaison combinatoire probabiliste et statistique, à savoir les conditions de logement et la mobilité urbaine. Les résultats ont mis en évidence que l'égalité des performances statistiques observées n'est pas forcément la conséquence logique ou immédiate de l'équiprobabilité individuelle de sélection. Par ailleurs, les écarts enregistrés sont tributaires des probabilités d'inclusion, de la combinatoire et de l'estimateur de Horvitz-Thompson.

Mots-clés : échantillonnage aléatoire, échantillonnage systématique, combinatoire, probabilités et structure spatiale urbaine

Digital Object Identifier (DOI): <https://doi.org/10.5281/zenodo.23640479>

1. Introduction

La ville de Kinshasa est confrontée depuis plusieurs années à une forte pression démographique qui entraîne des transformations importantes, souvent spontanées, de son organisation et de son aménagement spatial (Lelo Nzuzi, 2008) . Les conditions de logement et les pratiques de mobilité urbaine en constituent des manifestations essentiellement significatives. Ces transformations posent des défis importants non seulement aux décideurs politico-administratifs et aux aménageurs urbains, mais également aux populations concernées, dont les conditions de vie peuvent être affectées par les difficultés liées à l'habitat et aux déplacements. Devant cet état de choses, il est clair que la solution peut dépendre, d'une part de la valeur des décisions et des interventions urbaines et d'autre part de la définition et détermination de la population cible, de la taille de

l'échantillon et du choix de méthodes d'échantillonnage rattachées aux structures spatiales urbaines.

Dans ce contexte, les travaux de Wu et Thompson (2020) ; Steven K. Thompson (2012) et D.R Bellhouse (2014) ont démontré que les mathématiques, dans sa sous- branche statistique, offrent des ingrédients scientifiques permettant la collecte des données qui conduisent à des résultats adéquats. Parmi ces ingrédients, les méthodes d'échantillonnages aléatoire simple et systématique sont considérés comme des outils mathématiques pertinents et efficaces.

Cependant, dans les enquêtes consacrées aux structures spatiales urbaines dans le contexte de l'Institut supérieur d'architecture et d'urbanisme de Kinshasa (2022, 2023,2024 et 2025), en République démocratique du Congo, le choix de l'échantillonnage aléatoire simple basée sur les comparaisons traditionnelles des moyennes, occupe une place importante par rapport à l'échantillonnage systématique. Or, lorsque les deux méthodes d'échantillonnage présentent une équiprobabilité individuelle d'inclusion ; peuvent conduire à des estimations ponctuelles identiques ou proches. Ainsi, existe-t-il un risque de considérer qu'ils possèdent nécessairement les mêmes propriétés statistiques et les mêmes performances.

Dans l'optique de renforcer la rigueur scientifique des comparaisons portant sur l'organisation spatiale des populations urbaines, il apparaît nécessaire de dépasser la seule comparaison des estimations ponctuelles et d'intégrer les dimensions combinatoire et probabiliste ainsi que les propriétés des estimateurs, notamment l'estimateur de Horvitz-Thompson, les probabilités d'inclusion individuelles et conjointes, la variance et la précision. Cette préoccupation conduit à la question de savoir en quoi ces différentes dimensions permettent de caractériser les différences entre l'échantillonnage aléatoire simple et l'échantillonnage systématique dans les enquêtes relatives aux structures spatiales urbaines, notamment celles portant sur l'habitat et la mobilité urbaine.

Sachant que les échantillonnages aléatoire simple et systématique peuvent présenter une équiprobabilité individuelle d'inclusion, mais leurs performances peuvent néanmoins différer en matière de précision et de variance des estimateurs, en raison de leurs structures combinatoires et de leurs probabilités d'inclusion conjointes. Ces performances varient également selon les

caractéristiques des structures spatiales urbaines étudiées, notamment celles relatives à l'habitat et à la mobilité urbaine.

La présente étude consiste à comparer les propriétés combinatoires, probabilistes et statistiques de deux plans d'échantillonnage dans les enquêtes urbaines. Son importance est à la fois théorique, méthodologique et pratique, puisqu'elle met en relation la combinatoire, les probabilités d'inclusion, l'estimation statistique et la structure spatiale des phénomènes urbains.

Le choix porté sur ce sujet se justifie par la nécessité de disposer d'une approche intégrée permettant de mieux choisir le plan d'échantillonnage adapté aux enquêtes urbaines. En associant combinatoire, probabilités, estimation statistique et analyse spatiale, cette recherche cherche ainsi à contribuer à une meilleure conception des enquêtes sur l'habitat et la mobilité urbaine.

2. Revue critique de la littérature et formalismes mathématiques mis en œuvre

2.1. Revue critique de la littérature

Les recherches relatives aux méthodes d'échantillonnage associées aux fondements mathématiques ont amélioré la qualité des décisions de plusieurs travaux. La contribution de Horvitz et Thompson (1952), en matière des méthodes d'échantillonnages, a établi un cadre général d'estimation fondé sur les probabilités d'inclusion, donnant naissance à l'estimateur de Horvitz-Thompson, qui permet d'estimer sans biais un total de population à partir des unités sélectionnées et de leurs probabilités d'inclusion.

Plus tard, Särndal, Swensson et Wretman (1992) ont ensuite développé une approche générale de l'échantillonnage et de l'estimation, en privilégiant les probabilités d'inclusion, à l'estimation de la variance et à la comparaison des plans d'échantillonnage. Leur contribution montre que l'évaluation d'un estimateur ne peut être limitée à sa valeur ponctuelle, mais doit également considérer sa précision et sa variance.

Dans cette perspective, Thompson (2012) différencie clairement l'échantillonnage aléatoire simple de l'échantillonnage systématique et souligne l'importance de la structure de la méthode dans la conception des enquêtes. Il présente l'échantillonnage systématique comme une méthode reposant sur une organisation régulière des unités et consacre également une place importante à l'échantillonnage spatial, ce qui rejoint directement l'orientation de la présente recherche. Dans le même d'ordre, Bellhouse (2014) renforce, quant à lui, l'intérêt pratique de

l'échantillonnage systématique en raison de sa simplicité et de sa facilité de mise en œuvre. Toutefois, son caractère régulier distingue fondamentalement la méthode d'échantillonnage aléatoire simple et justifie l'examen de ses propriétés selon la structure de la population étudiée.

Cependant, ces travaux ont, dans une certaine mesure, abordé chacun en ce qui le concerne, les différences entre les méthodes d'échantillonnages aléatoire et systématique. Ils les abordent dans le sens de comparaison en stigmatisant le caractère important de l'une ou l'autre méthode. La présente étude cherche à rapprocher ces dimensions dans un même cadre comparatif, en examinant simultanément la structure combinatoire, les probabilités individuelles et conjointes, l'estimateur de Horvitz-Thompson, la variance et les caractéristiques spatiales. L'originalité recherchée réside ainsi dans l'application de cette approche intégrée aux structures urbaines de l'habitat et de la mobilité.

Alors que les travaux antérieurs ont largement mobilisé les notions de biais, de variance et de précision des estimateurs pour évaluer les performances des plans d'échantillonnage, la présente étude montre que l'intégration de la combinatoire, considérée comme une structure mathématique de l'espace des échantillons possibles, des probabilités d'inclusion individuelles et conjointes, ainsi que des propriétés des estimateurs, permet d'approfondir la caractérisation et l'interprétation des différences entre l'échantillonnage aléatoire simple et l'échantillonnage systématique.

2.2. Formalismes mathématiques mis en œuvre

2.2.1 Population, échantillon et échantillonnage

En référence aux définitions avancées par Bellhouse (2014) ; Wu, C. & Thompson, M (2020) et RAHIM Bilal (2018), une **population** E , de taille N , désigne l'ensemble fini des unités statistiques sur lesquelles porte une étude. Une unité statistique peut correspondre à un individu, un ménage, une entreprise ou toute autre entité d'observation. En d'autres termes, si l'on a une population finie $U = \{1, 2, \dots, N\}$ de taille N , dans laquelle on souhaite sélectionner un échantillon de taille n . On note : y_i : valeur de la variable étudiée pour l'unité i ; I_i : indicatrice d'inclusion de l'unité i ; p_i : probabilité individuelle d'inclusion de l'unité i ; p_{ij} : probabilité que les unités i et j soient incluses simultanément.

2.2.2 Méthodes d'échantillonnage aléatoire simple et échantillonnage systématique

L'échantillonnage aléatoire simple (EAS), considérée comme méthode probabiliste, est une méthode de sélection dont chaque sous-ensemble possible de taille n possède la même probabilité d'être sélectionné. Dans l'échantillonnage systématique (ES), les unités sont sélectionnées à intervalles réguliers caractérisant le pas ($k = N/n$). Le pas d'échantillonnage est défini comme un nombre k d'éléments qu'on saute entre deux individus sélectionnés dans la population N . Il indique l'intervalle régulier entre les individus sélectionnés.

2.2.3 Espace et structure combinatoire

Dans l'échantillonnage aléatoire simple sans remise, l'espace des échantillons possibles est :

$$\Omega = \{s \subset U : |s| = n\} \text{ et le nombre total d'échantillons possibles est : } |\Omega| = \binom{N}{n} = \frac{N!}{n!(N-n)!}.$$

2.2.4. Fondements probabilistes

➤ Probabilités de sélection

Dans l'échantillonnage aléatoire simple, chaque échantillon possède la même probabilité : $P(s) = 1/C(N, n)$. Par contre dans le systématique, lorsque $k = N/n$ et que le départ est uniformément choisi parmi les k possibilités : $P(s_r) = 1/k = n/N$ pour chacun des échantillons systématiques possibles.

➤ Probabilité individuelle d'inclusion

L'indicatrice d'inclusion est définie par : $I_i = 1$ si i appartient à s et : $I_i = 0$ sinon.

La probabilité individuelle d'inclusion est : $p_i = P(I_i = 1)$

➤ Probabilités d'inclusion conjointe

Pour deux unités distinctes i et j : les probabilités d'inclusion conjointes sont définies par $p_{ij} = P(I_i = 1 \text{ et } I_j = 1)$.

Dans le cas aléatoire simple : $p_{ij} = n(n-1) / [N(N-1)]$. On peut également écrire :

$$p_{ij} = [n/N] \times [(n-1)/(N-1)].$$

Pour ce qui concerne le systématique la probabilité dépend de la position des unités i et j

dans la liste. Dans le cas classique où elles peuvent être sélectionnées ensemble : $p_{ij} = 1/k = n/N$ et lorsqu'elles ne peuvent pas appartenir au même échantillon : $p_{ij} = 0$.

3. METHODOLOGIE DU TRAVAIL

La méthodologie de cette étude s'inscrit dans le cadre de l'approche quantitative, comparative et déductive, fondée sur les principes de la théorie des sondages probabilistes et de l'analyse combinatoire. Elle vise la comparaison des propriétés mathématiques et statistiques d'échantillonnages aléatoire simple et systématique, à travers deux applications numériques portant sur les conditions de logement et la mobilité urbaine.

L'étude s'appuie sur deux populations finies représentatives de problématiques urbaines. La première concerne une population de 200 ménages, dont un échantillon de 50 ménages est sélectionné afin d'évaluer les conditions de logement. De ces 50 ménages, 35 représentent la condition de logement étudiée. La seconde porte sur une population de 250 000 habitants, dans laquelle un échantillon de 125 personnes est constitué pour analyser les comportements de mobilité selon quatre modes de déplacement : la moto, le taxi, la marche à pied et le taxi-bus, avec respectivement les nombres de 33 ,26,41 et 25.

Sur le plan méthodologique, la comparaison de ces deux méthodes ne s'est pas limitée aux estimations des paramètres de la population. En vue de comprendre les mécanismes qui gouvernent les performances des plans d'échantillonnage probabilistes, plusieurs dimensions complémentaires ont été mis à profit, notamment la structure combinatoire des plans d'échantillonnage, les probabilités de sélection, les probabilités d'inclusion de premier et de second ordre, les estimateurs de la moyenne et du total, l'estimateur de Horvitz-Thompson, les variances, le biais, l'erreur quadratique moyenne ainsi que l'efficacité relative des plans de sondage.

L'interprétation des résultats a été réalisée selon une démarche comparative articulée autour de quatre axes : la comparaison des structures combinatoires, l'analyse des probabilités d'inclusion, l'évaluation des performances des estimateurs et l'examen des implications des résultats pour les enquêtes sur les conditions de logement et la mobilité urbaine.

4. Résultats

4.1. Présentation de la Situation n°1 : Habitat de conditions de logement

Cette situation est relative à un échantillon de 50 ménages issu d'une population de 200 ménages. Y défini par $Y_i = \sum_{i=0}^1$ qui représente les 7 indicateurs de conditions de logement, à savoir : qualité du bâtiment, accès à l'eau, assainissement, électricité, aération, nombre de pièces et sécurité. Si $y = 0$, alors les conditions de logement sont funestes et dans le cas où les conditions de logement sont adéquates ou non funestes, alors $Y=1$.

4.2 Résultats synthétiques de deux méthodes de la situation n°1

Indicateur	Formalisme	Aléatoire simple	Systématique
Population	N	200	200
Echantillon	n	50	50
Échantillon des logements défavorables	-	35	35
Échantillon des logements non défavorables	--	15	15
Proportion des logements défavorables	$\frac{\text{Effectif observé}}{n}$	70%	70%
Intervalle systématique	$\frac{N}{n}$		4
Nombre de départs systématiques possibles	$\frac{N}{n}$	-	4
Structure combinatoire	$\binom{N}{n} = \frac{N!}{n!(N-n)!}$	$4,54.10^{47}$	10^{608}
Probabilité	$P(S) = \frac{1}{\binom{N}{n}}$	$2,20.10^{-48}$	25 %
Probabilité individuelle(de sélection)	$\pi_i = \frac{n}{N}$	25 %	25 %
Probabilité d'inclusion conjointe	$\pi_{ij} = \frac{n(n-1)}{N(N-1)}$	6,15578%	0 ou 0,25
Poids Horvitz-Thompson	$\frac{1}{\pi_i}$	4	4

Pas	$K = \frac{N}{n}$	-	4
Total Horvitz-Thompson défavorable	$T_{HT} = \sum_{pi} \frac{y_i}{p_i}$, $i \in s$	140 ménages	140 ménages
Proportion Horvitz-Thompson	$Prop HT = \frac{Total\ estimé}{N}$	70%	70%
Estimateur de la moyenne	$\tilde{Y} = \frac{1}{n} \sum_{i \in s} y_i$	$5,48 \cong 5$	5,5
Variance de l'estimateur de la moyenne	$V(\hat{Y}) = (1 - \frac{n}{N}) \frac{S^2}{n}$	0,003214	0,000200
Ecart type	$E = \sqrt{V(\hat{Y})} = (1 - \frac{n}{N}) \frac{S^2}{n}$	0,05669	0,01414
Estimateur de Horvitz-Thomson	$EHT = \text{Cas observés} \times \frac{1}{\pi_i}$	140 ménages	140 ménages
Proportion Horvitz-Thompson	Proportion estimée $E = \frac{HT}{N}$	70 %	70 %
Biais	$B(\hat{Y}) = E(\hat{Y}) - \tilde{Y}$	0	0
Efficacité relative	$ER = \frac{V(\tilde{Y})_{EAS}}{\sqrt{V(\hat{Y})}_{ES}}$	1	16,07
Ecart Quadratique Moyen(EQM)	$EQM = \text{Variance-biais}$	0,003214	0,0002000

Tableau n°1: Résultats synthétiques de deux méthodes d'échantillonnages de l'habitat

Le tableau n°1 montre que les deux méthodes d'échantillonnage possèdent deux structures combinatoires différentes, avec celle de l'échantillonnage aléatoire est extrêmement riche. Aussi, il révèle l'équiprobabilité et la même estimation ponctuelle de la proportion des ménages présentant la condition de logement (70 %) et au même estimateur de Horvitz-Thompson du total (140 ménages). Cette similitude provient de l'égalité des probabilités individuelles d'inclusion ($p_i = 0,25$) et du même poids d'extrapolation (4).

Cependant, les performances statistiques diffèrent au niveau de la précision. L'échantillonnage systématique présente une variance plus faible (0,000200 contre 0,003214) et une efficacité

relative de 1,144, indiquant un gain de précision d'environ 16,07 % par rapport à l'échantillonnage aléatoire simple.

4.3. Présentation de la situation n° 2 : Cas de la mobilité urbaine

La population étudiée comprend 250 000 habitants, dont un échantillon de 125 personnes est sélectionné.

4.4. Résultats synthétiques de deux méthodes de la situation n°2

Indicateur	Formalisme		Moto	Taxi	Marche	Taxi-bus
Population	N		250000	250000	250000	250000
Échantillon	n		125	125	125	125
Effectif observé	-		33	26	41	25
Proportion observée	$\frac{\text{Effectif observé}}{n}$		26,40%	20,80%	32,80%	20,00%
K = Intervalle systématique	$\frac{N}{n}$		2000	2000	2000	2000
Nombre de séquences ES possibles	$\frac{N}{n}$		2000	2000	2000	2000
Structure combinatoire	$\binom{N}{n} = \frac{N!}{n!(N-n)!}$	EAS	$1,20 \cdot 10^{47}$	$1,20 \cdot 10^{47}$	$1,20 \cdot 10^{47}$	$1,20 \cdot 10^{47}$
		ES				
Probabilité individuelle (de sélection) P_i	$\pi_i = \frac{n}{N}$		0,0005	0,0005	0,0005	0,0005
Poids Horvitz-Thompson	$1/\pi_i$		2000	2000	2000	2000
	$\pi_{ij} = \frac{n(n-1)}{N(N-1)}$	EAS	$2,480 \cdot 10^{-7}$	$2,480 \cdot 10^{-7}$	$2,480 \cdot 10^{-7}$	$2,480 \cdot 10^{-7}$

Probabilité d'inclusion conjointe	$\pi_{ij} = \frac{1}{k} = \frac{n}{k}$	ES	0,0005 ou 0	0,0005 ou 0	0,0005 ou 0	0,0005 ou 0
Estimateur de Horvitz-Thomson Total	$T_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i}$		66000	52000	82000	50000
Proportion Horvitz-Thompson	Prop HT = $\frac{\text{Total estimé}}{N}$		26,40%	20,80%	32,80%	20,00%
Estimateur de la moyenne	$\tilde{Y} = \frac{1}{n} \sum_{i \in s} y_i$		0,2640	0,2080	0,3280	0,2000
Variance de l'estimateur de la moyenne	$V(\hat{Y}) = (1 - \frac{n}{N}) \frac{S^2}{n}$	EAS	0,001554	0,001317	0,001762	0,001279
		ES	0,001602	0,001333	0,001758	0,001179
Ecart type	$\sqrt{V(\hat{Y})} = (1 - \frac{n}{N}) \frac{S}{\sqrt{n}}$	EAS	0,03942	0,03629	0,04198	0,03577
		ES	0,04002	0,03651	0,04193	0,03434
Biais Thompson-Horvitz	$B(\hat{Y}) = E(\hat{Y}) - \bar{y}$		0	0	0	0
Efficacité relative	$ER = \frac{V(\hat{Y}_{EAS})}{V(\hat{Y}_{SYS})}$		0,970	0,988	1,002	1,085
Ecart Quadratique Moyen (EQM)	EQM = Variance - biais ²	EAS	0,001554	0,001317	0,001762	0,001279
	EQM = Variance - biais ²	ES	0,001602	0,001333	0,001758	0,001179

Tableau n°2 : Résultats synthétiques de deux méthodes d'échantillonnages de la mobilité

Les résultats présentés dans le tableau n°2 montrent que, dans la mobilité urbaine, les quatre

modes de déplacement, bien que sélectionnées suivant différentes proportions observées avec des estimateurs de la moyenne proche, ont la même probabilité individuelle d'inclusion de 0,0005, soit 0,5 % et le même poids HT de 2000. Par ailleurs, il est observé que la variance diffère selon la proportion étudiée. La marche, avec une proportion de 32,8 %, possède la variance la plus élevée (0,001762 dans EAS et 0,001758 dans ES, car elle est la modalité la plus éloignée des extrêmes (0 ou 1) et présente donc la plus grande variabilité. Dans l'entre temps, pendant que le tableau présente les mêmes proportions de Horvitz-Thomson de l'ordre de 70 % avec les mêmes biais ; et de l'autre côté les efficacités relatives et les EQM demeurent différents.

5. Discussion

les résultats obtenus dans les deux situations étudiées, notamment les conditions de logement et la mobilité urbaine, montrent l'équiprobabilité individuelle d'inclusion y relative, bien que les mécanismes de constitution des échantillons restent différents.

En effet, dans la situation des conditions de logement, avec un échantillon de 50 ménages sur 200 ménages, la proportion observée de ménages présentant une condition défavorable est 70 %. La proportion de ménages présentant une condition favorable est de 30 %. Pour les deux méthodes, chaque ménage possède la même probabilité individuelle d'être sélectionné de 0,25. L'estimateur de Horvitz-Thompson du nombre total de ménages présentant une condition défavorable devient alors 140 ménages. La proportion estimée pour l'ensemble de la population est donc de 70 %. Dans la configuration systématique construite de 200 ménages les quatre proportions correspondantes sont donc : 70 %, 68 %, 72 % et 70. Bien qu'il existe une équiprobabilité individuelle d'inclusion de 25 %, les résultats obtenus sont différents.

Il en est de même pour la situation de mobilité urbaine qui constitue le deuxième axe de discussion, où 125 personnes sont sélectionnées sur les 250 000 habitants, les deux méthodes présentent une probabilité individuelle d'inclusion de 0,05 % et un poids de Horvitz-Thompson de 2 000. Pourtant, les proportions estimées diffèrent selon les modes : 26,40 % de moto contre 20,80 % de taxi ; 32,80 % de marche contre 20 % de taxi-bus.

Avec le pas systématique de 2 000, il existe donc 2 000 positions possibles de départ dans le plan systématique. Chaque personne possède une probabilité individuelle de sélection égale à 0,05 %. Le poids associé à chaque personne sélectionnée est donc : 2 000. Les estimations de

Horvitz-Thompson des effectifs deviennent : Moto (66 000), Taxi (52 000), Marche à pied (82 000) et Taxi-bus (50 000).

Peut-on croire que les écarts observés de probabilité permettent de confirmer la supériorité ou la performance d'une méthode. Nous pensons qu'ils montrent plutôt que l'équiprobabilité individuelle n'implique pas l'identité des échantillons ni celle des estimations obtenues. En d'autres termes, cette propriété commune ne signifie pas que les deux méthodes produisent nécessairement les mêmes échantillons, les mêmes structures de dépendance entre unités, ni les mêmes niveaux de variance et de précision. L'avantage provient plutôt de la manière dont les unités sont réparties dans les différentes séquences du plan systématique.

Ce constat rejoint les travaux de Bethlehem (2009), Lohr (2021) et Kim (2024), qui montrent que le choix du plan de sondage doit être étudié conjointement avec les propriétés des estimateurs et de leur variance. Il rejoint également les travaux classiques de Madow (1946) et Cochran (1946), qui montrent que les performances du sondage systématique dépendent de l'organisation de la population et de l'ordre dans lequel les unités sont disposées.

En référence aux valeurs de variance, dans le cas d'habitat, pour l'EAS, la variance estimée de la proportion est d'environ 0,003214 avec un écart-type d'environ 0,05669, soit environ 5,67 %.

La variance obtenue pour l'ES est alors : 0,000200 avec un écart-type de 0,01414, soit environ 1,41 %. L'efficacité relative obtenue est de 16,07. Autrement dit, dans cette configuration particulière, la variance obtenue avec l'EAS est environ 16 fois supérieure à celle obtenue avec l'ES. Ce constat montre que l'échantillonnage aléatoire simple offre un espace extrêmement vaste de configurations possibles, alors que le systématique sélectionne les ménages selon une structure régulière. Les résultats de variance obtenus à partir des proportions observées donnent des valeurs différentes de variances et écart-types associées aux différents modes de déplacement, selon chaque méthodes d'échantillonnage. Le biais empirique demeure toutefois pratiquement nul dans les deux plans, ce qui montre que la différence observée relève principalement de la variabilité des estimations plutôt que d'un biais systématique.

La comparaison dans le cas de la mobilité donne une même efficacité dans la configuration

EAS et des différences dans le cas de l'ES : 0,970 pour la Moto, 0,988 pour le Taxi, 1,002 pour la Marche à pied et 1,085 pour le Taxi-bus. Contrairement au cas de l'habitat, l'ES ne présente pas ici un avantage massif pour toutes les catégories. Son efficacité est légèrement inférieure à celle de l'EAS pour la moto et le taxi, pratiquement identique pour la marche à pied, et supérieure pour le taxi-bus. Cela montre que l'efficacité d'une méthode d'échantillonnage n'est pas une propriété absolue de la méthode elle-même. Cette variation montre que la performance d'un plan d'échantillonnage dépend non seulement de ses propriétés probabilistes générales, mais également de la structure du phénomène étudié et de son organisation dans la population.

En accord avec les travaux de Horvitz et Thompson (1952) qui fournissent le fondement de l'utilisation des probabilités d'inclusion dans l'estimation ; il est vrai le poids dépend de la probabilité individuelle d'inclusion, alors que la variance dépend également de la structure conjointe des inclusions. Cela ne confirme pas le fait que les poids HT soient identiques pour les deux plans à taille d'échantillon entraîne que leurs performances doivent être identiques.

Etant donné que les recherches de Särndal, Swensson et Wretman (1992) montrent l'importance de relier l'estimation et la variance au plan d'échantillonnage et considérant l'idée de Singh et Singh (1977) axée sur des procédures d'estimation de variance adaptées au systématique et sur la comparaison de son efficacité avec celle du tirage aléatoire simple, il est évident de dire que l'efficacité relative du systématique et de l'aléatoire simple peut varier selon la structure de la population et le plan considéré. Aussi, il convient également de souligner que l'estimation de la variance du systématique ne doit pas être traitée automatiquement comme celle de l'aléatoire simple et la structure particulière de la méthode doit être prise en compte.

Ces résultats sont cohérents avec le raisonnement de Horvitz et Thompson (1952), selon lequel l'estimation par sondage probabiliste doit tenir compte des probabilités d'inclusion, mais ils conduisent à approfondir cette perspective en considérant également les probabilités d'inclusion conjointes. En effet, l'égalité des probabilités individuelles p_i entre les deux méthodes n'implique pas l'égalité des probabilités conjointes π_{ij} . Or, celles-ci interviennent directement dans la

variance de l'estimateur de Horvitz-Thompson. C'est donc la structure des dépendances entre les unités sélectionnées qui contribue à expliquer les différences de précision observées.

Contrairement au cas de l'habitat, l'échantillonnage aléatoire simple ne présente pas ici un avantage massif pour toutes les catégories. Son efficacité est légèrement inférieure à celle du systématique pour la moto et le taxi, pratiquement identique pour la marche à pied, et supérieure pour le taxi-bus. Cela montre que l'efficacité d'une méthode d'échantillonnage n'est pas une propriété absolue de la méthode elle-même. Elle dépend de la relation entre la méthode de sondage, l'organisation de la population et la variable étudiée.

À l'inverse à l'habitat, qui possède une certaine stabilité spatiale, la mobilité correspond à des déplacements, des flux, des interactions et des relations entre différents lieux. Les travaux de Pumain (2010) sur l'analyse spatiale mettent justement l'accent sur les distributions spatiales, les réseaux, les flux et les interactions qui structurent les systèmes urbains. Cette conception permet de comprendre pourquoi un même plan de sondage peut avoir des performances différentes lorsqu'il est appliqué à des ménages localisés dans l'espace ou à des personnes caractérisées par leurs déplacements. Dans cette perspective, les travaux de Roncayolo (2002) permettent également de considérer la ville comme une organisation spatiale évolutive dans laquelle les formes urbaines, les usages et les pratiques se transforment dans le temps. L'échantillonnage de la mobilité doit donc être pensé non seulement comme une opération statistique, mais également comme une opération adaptée à une structure spatiale et dynamique. Les résultats obtenus apportent ainsi un argument supplémentaire à l'hypothèse générale de l'étude : deux méthodes d'échantillonnage peuvent être équiprobables au niveau individuel tout en présentant des performances différentes au niveau global. Les tableaux n°1 et 2 montrent également que ces différences ne sont pas uniformes : elles varient selon la structure spatiale considérée, ce qui est particulièrement pertinent pour les enquêtes urbaines portant sur l'habitat et la mobilité.

En conséquence, la comparaison des échantillonnages aléatoire et systématique ne devrait pas se limiter aux estimations ponctuelles ou aux probabilités individuelles d'inclusion. Elle doit intégrer simultanément la structure combinatoire de l'espace des échantillons, les probabilités

d'inclusion conjointes, les propriétés de l'estimateur de Horvitz-Thompson, la variance, le biais, l'erreur quadratique moyenne et l'efficacité relative.

CONCLUSION

Cette recherche a pour objectif de comparer le combinatoire et probabilités des échantillonnages aléatoire simple et l'échantillonnage systématique à travers deux applications portant sur les conditions de logement et la mobilité urbaine. Les analyses réalisées ont montré que ces deux méthodes reposent sur les mêmes principes probabilistes de base, mais qu'elles présentent des structures combinatoires et des propriétés statistiques distinctes qui influencent directement la précision des estimations. Aussi, elles confirment que les deux méthodes d'échantillonnage offrent aux unités de la population des probabilités individuelles d'inclusion identiques et permettent d'obtenir des estimateurs non biaisés, notamment grâce à l'estimateur de Horvitz-Thompson ; et ne suffit pas à garantir une efficacité statistique équivalente.

Les différences observées proviennent principalement des probabilités d'inclusion de second ordre, de l'organisation de la base de sondage et de la structure combinatoire propre à chaque plan. En d'autres termes, il est démontré les limites de l'équiprobabilité individuelle d'inclusion dans la caractérisation des performances de deux méthodes d'échantillonnage, celles-ci pouvant différer selon leur structure de sélection, les configurations spatiales considérées et les propriétés de leurs estimateurs.

L'étude montre que l'échantillonnage aléatoire simple offre une plus grande richesse combinatoire et une meilleure robustesse théorique, tandis que l'échantillonnage systématique présente des avantages opérationnels importants, notamment en matière de simplicité de mise en œuvre, de réduction des coûts et de couverture spatiale des unités enquêtées. Le choix entre ces deux méthodes doit donc être effectué en fonction des caractéristiques de la population, des objectifs de l'enquête et des contraintes pratiques de collecte des données.

Cette étude contribue ainsi à renforcer la rigueur des comparaisons entre plans d'échantillonnage et offre des orientations pour la conception et l'analyse des enquêtes quantitatives appliquées aux sciences urbaines, notamment dans les domaines de l'habitat et de la mobilité. Néanmoins, l'étude se limite à la comparaison de deux méthodes d'échantillonnage dans des situations

simplifiées. Ces limites ouvrent la voie à des investigations axées sur des données plus large confrontées à d'autres méthodes d'échantillonnage adaptées à la diversité des structures spatiales urbaines. De telles perspectives permettraient de vérifier empiriquement les résultats obtenus et d'élargir la portée de l'approche combinatoire et probabiliste proposée.

Bibliographie

- [1]. Baily, A. S., & Béguin, H. (2018). *Introduction à la géographie quantitative*. Paris : Armand Colin ;
- [2]. Bellhouse, D. R. (2014). *A brief history of random sampling*. In *Handbook of Statistics: Sample Surveys: Design, Methods and Applications Elsevier*;
- [3]. Bethlehem, J. (2009). *Applied survey methods: A statistical perspective*. Hoboken: Wiley;
- [4]. Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons;
- [5]. Fuller, W. A. (2011). *Sampling Statistics*. Hoboken: John Wiley & Sons;
- [6]. Horvitz, D. G., & Thompson, D. J. (1952). *A generalization of sampling without replacement from a finite universe*. *Journal of the American Statistical Association*, 47(260), 663–685.
- [7]. Kim, J. K. (2024). *Statistical Foundations of Survey Sampling and Inference*. Springer International Publishing .
- [8]. Lelo Nzuzi, F (2008). *Kinshasa : Planification urbaine et aménagement*. Editions L'Harmattan.
- [9]. Lohr, S. L. (2021). *Sampling: Design and Analysis* (3rd ed.). Boca Raton: CRC Press;
- [10]. Pumain, D., & Saint-Julien, T. (2010). *L'analyse spatiale : localisations dans l'espace*. Paris : Armand Colin ;
- [11]. Roncayolo, M. (2002). *Lectures de la ville : formes et transformations urbaines*. Marseille : Éditions Parenthèses ;
- [12]. Särndal, C. E., Swensson, B., & Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag;
- [13]. Tillé, Y. (2020). *Sampling Algorithms*. Cham: Springer International Publishing;
- [14]. Thompson, S. K. (2012). *Sampling* (3rd ed.). Hoboken: John Wiley & Sons;
- [15]. Wu, C., & Thompson, M. E. (2020). *Sampling Theory and Practice*. Cham: Springer International Publishing;
- [16]. Yates, F. (1981). *Sampling Methods for Censuses and Surveys*. London: Griffin;