



7e colloque francophone sur les Sondages 2012

Analyse des dépenses de consommations des ménages de la ville de Ouagadougou : application des techniques de rééchantillonnage pour l'estimation de la variance

Présenté par
OUEDRAOGO Wendingoudi Tidiane

Plan de la présentation

- Introduction : contexte et justification
- Objectif
- Données
- Méthodologie
- Principaux résultats

Introduction

- Dans les pays en voie de développement, la plupart des grosses opérations d'enquêtes par sondage se fait par un échantillonnage (souvent stratifié) à plusieurs degrés
- Présente de nombreux avantages :
 - Réduction du coût et temps de la collecte
 - Économie de conception d'une liste exhaustive de toute la population
- Cependant, la complexité de l'échantillonnage à plusieurs degrés rend complexe le calcul des estimations de la variance (ou précision)
- Encore plus complexe pour les estimateurs non linéaires

Introduction (suite)

- Trouver d'autres méthodes d'estimation de la variance
- Dans la littérature, plusieurs méthodes existent pour estimer la variance mais les plus importantes
 - Les techniques de linéarisation
 - Les techniques de rééchantillonnage

Objectifs

- L'objectif principal : appliquer des techniques de ré-échantillonnage
- Pour l'estimation des variations des dépenses de consommations des ménages de la ville de Ouagadougou
- Spécifiquement : l'analyse portera sur l'estimation de
 - Des dépenses de consommation moyenne
 - Les quintiles des dépenses de consommation

Les données

Elles proviennent de l'enquête sur les dépenses de consommation des ménages de la ville de Ouagadougou en 2008.

Cette enquête s'était inscrite dans le cadre du projet de rénovation de l'indice harmonisé des prix à la consommation (IHPC) que pilotent AFRISTAT et les instituts nationaux de statistiques.

Les données (suite)

- Le plan de sondage est un sondage aréolaire à deux degrés.
- Au premier degré, il a été tiré 84 unités primaires ou ZD (zones de dénombrement) à probabilité égale
- Au second degré 12 ménages ont été tirés à probabilité égale.

Méthodologie

- Basée sur les techniques de rééchantillonnage
- L'idée de base repose sur la construction d'un nombre R d'échantillons et à déterminer la statistique d'intérêt pour chaque nouvel échantillon
- Les principales méthodes : le jackknife, les demi-échantillons équilibrés et le bootstrap
- le jackknife et le bootstrap sont utilisés dans cette présentation

Méthodologie (suite)

Le jackknife

➤ Le principe de base de cette méthode repose sur la suppression de groupe d'observations et à ré estimer la statistique d'intérêt avec chacune des observations restantes. Soit θ le paramètre à estimer, $\hat{\theta}_{-i}$ est l'estimation du paramètre (en supprimant l'observation i), n étant la taille de l'échantillon la variance est alors estimée par :

$$\hat{V}_{jk}(\hat{\theta}) = \frac{n-1}{n} \sum_{i=1}^n (\hat{\theta}_{-i} - \hat{\theta})^2$$

➤ Dans le cas du sondage 2 degrés, l'application du jackknife consiste à supprimer chaque unité primaire avec toutes ses observations et à calculer la statistique sur le reste des unités primaires.

Méthodologie (suite)

Le bootstrap

➤ il consiste à reproduire le mécanisme d'échantillonnage d'origine qui a engendré les données et recalcule la statistique avec les échantillons tirés aléatoirement de l'échantillon de départ.

➤ Si B est le nombre de réplication alors la variance bootstrap de $\hat{\theta}$ est donné par

➤
$$\widehat{V}_{Boot}(\hat{\theta}) = (B - 1)^{-1} \sum_{b=1}^B (\hat{\theta}_b - \bar{\hat{\theta}})^2 \quad \text{où} \quad \bar{\hat{\theta}} = B^{-1} \sum_{b=1}^B \hat{\theta}_b$$

Méthodologie (suite)

Le bootstrap (suite)

- Dans le cas d'un sondage à plusieurs degrés (où à chaque degré un sondage aléatoire simple est effectué), la méthode de bootstrap appliqué est de rééchantillonner les unités primaires.

Les principaux résultats

- Estimation de la consommation moyenne par ménage

	Estimation (FCFA)	bootstrap		Jackknife	
		écart type	Coef. variation (%)	écart type	Coef. variation (%)
moyenne	2106101	137302	6,52	130398	6,19

- la qualité des données relative à la variable d'intérêt est bonne pour les deux types d'estimation de la variance
- Ainsi, à Ouagadougou, les ménages consomment en moyenne 2 106 101Fcfa par an soit environ 5 780Fcfa par jour
- Cependant cela cache une certaine disparité au niveau des citoyens de Ouagadougou.

Les principaux résultats (suite)

➤ Estimation des quintiles de dépenses de consommation

	Estimation (FCFA)	bootstrap		Jackknife	
		écart type	Coef variation (%)	écart type	Coef variation (%)
1er Quintile	711959,1	57474,74	8,07	55270,1	7,76
2ème Quintile	1145100,8	75134,88	6,56	75452,22	6,59
3ème Quintile	1781536,4	95037,11	5,33	94251,83	5,29
4ème Quintile	2916691,8	211337,49	7,25	195822,28	6,71
5ème Quintile	21718240,6	1039616,15	4,79	1035552,73	4,77

- 20% des ménages consomment moins 711 959Fcfa par an soit moins de 1950Fcfa par jour pour tous leurs besoins (y compris pour se loger, nourrir, s'habiller, etc.).
- Forte disparité entre les ménages, excepté le dernier quintile

Conclusion

- La mise en œuvre des techniques de réplication a permis d'évaluer la variance des certaines estimations dans le cadre d'un sondage aréolaire à 2 degrés.
- Cela permet ainsi d'apprécier la qualité des données relativement à la variable d'intérêt



Merci pour votre attention