Guide méthodologique sur l'échantillonnage et les estimations pour les ODD 2.4.1

Dramane Bako
Coordinateur de l'assistance technique
Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales
Division de la statistique de la FAO
Rome, Italie
Dramane.Bako@fao.org

Introduction

- L'indicateur 2.4.1 des objectifs de développement durable (ODD) est *la proportion des superficies agricoles exploitées de manière productive et durable*.
- La FAO à travers le programme « *Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales »* a réalisé des travaux méthodologiques sur le contenu et le calcul de l'indicateur
- Nous présentons ici une orientation méthodologique proposée aux pays en développement sur les questions d'échantillonnage

Définition

- Une exploitation agricole sera considérée comme durable si elle remplit certaines conditions liées aux thèmes ci dessous:

Dimension	N.	Thèmes
Économique	1	Productivité du sol
	2	Rentabilité agricole
	3	Résilience financière
Environnement	4	La santé du sol
	5	Santé de l'eau
	6	Biodiversité
Social	7	Travail décent
	8	Bien-être
	9	Accès à la terre

- Lorsqu'une exploitation est reconnue comme durable, toute sa superficie agricole est considérée comme durable
- L'ODD 2.4.1 est donc le rapport des superficies durables sur le total des superficies agricoles

Plan d'échantillonnage d'une enquête ad hoc

Population

- **Unités d'observation et Unités d'échantillonnage (PMRA, 2020)**
  - exploitations relevant du secteur des ménages, c'est-à-dire qui sont gérées par les ménages ("*exploitations familiales*") et
  - exploitations ne relevant pas du secteur des ménages, telles que les entreprises agricoles commerciales et les institutions gouvernementales («*exploitations non familiales*»)

Plan de sondage

- **Domaine d'estimation**: territoire national (pour la demande international)
- **Stratification**: toute typologie corrélée avec le maximum de thèmes d'intérêt de l'indicateur serait pertinente comme critère de stratification
- Plan de sondage (de GSARS, 2017)
  - Exploitations familiales: sondage à deux degré
  - Exploitations non familiales: sondage aléatoire simple

Taille de l’échantillon

- **Modélisation**
- $Y = \sum y_i$ = superficie agricole
- $Z = \sum z_i$ = superficie agricole durable
- $z_i = y_i$ si la superficie agricole de l'exploitation i est considéré comme durable; autrement $z_i = 0$
- p : proportion d'exploitations ayant une superficie agricole durable
- N_d nombre total d'exploitations dans le domaine d’estimation d
- L‘ODD 2.4.1 pour le domaine d est donc le ratio :

$$SDG2.4.1_d = \frac{\sum_1^{N_d} z_i}{\sum_1^{N_d} y_i} \quad (1)$$

Taille de l'échantillon

- $CV^2_{zU_d}$= carré du CV de Z
- taille d'échantillon minimale m_d

- $m_d = \frac{z^2_{\alpha/2} CV^2_{zU_d}}{\varepsilon^2}$ (2)
- $CV^2_{zU_d}$ peut être estimée à partir des données:
  - d'une enquête similaire précédente (quasi impossible pour la première collecte)
  - d'une enquête pilote (peut être non adapté/très cher)
  - une ancienne enquête agricole classique ayant simplement renseigné les superficies ou une enquête pilote sur les superficies (**plus réaliste mais comment?**)

Taille de l'échantillon

- Pour cette dernière approche nous devons mettre en relation les CV de Y et Z
- Considérons X suivant une Bernoulli $\mathcal{B}(p)$
- $Z = XY$
- $E(Z) = E(XY) = E(X)E(Y) = pE(Y)$

On montre que

- $V(Z) = pV(Y) + (p - p^2)E(Y)^2 \quad (3)$
- $CV_{zU_d}^2 = \frac{1}{p} CV_{yU_d}^2 + \frac{1-p}{p} \quad (5)$
- $m_d = \frac{z_{\alpha/2}^2}{\varepsilon^2} (\frac{1}{\hat{p}} CV_{yU_d}^2 + \frac{1-\hat{p}}{\hat{p}}) \quad (6)$
- $\hat{p}$ à partir d'avis d'experts, ou enquête pilote

Allocation de l'échantillon

- Neyman
- $m_{dh} = \widetilde{m}_d \frac{N_{dh} S_{zdh}}{\sum_h N_{dh} S_{zdh}} \quad (9)$
- On montre que
- $S^2_{zdh} = \hat{p} S^2_{ydh} + \hat{p}(1-\hat{p}) \bar{Y}^2_{dh} \quad (10)$

Cas spécifiques

- **Absence de base de sondage actualisée des exploitations:**
  - *exploitations familiales* : mettre à jour la liste des UPE (ZD, comtés, villages ...), sélectionner un échantillon d'UPE, établir une liste exhaustive d'exploitations dans les UPE échantillonnées pour la sélection finale
  - *exploitations non familiales* : utiliser les données administratives (par exemple, le registre des sociétés agricoles)

Cas spécifiques

- **Existence de base de sondage mais sans variable de stratification**
  - La stratification semble essentielle pour les besoins de l'indicateur
  - La recommandation est d'utiliser un échantillonnage en deux phases pour les besoins de stratification

Utilisation d’enquête existante

Enquêtes agricoles classiques avec l'exploitation agricole est l'unité d'échantillonnage

- Dans ce cas, il suffit de réajuster la taille de l'échantillon pour assurer une fiabilité de l'estimation de l'ODD 2.4.1
- Si n_d est la taille de l'échantillon actuel, on peut calculer la nouvelle taille réajustée m_d comme suit:

- $$m_d = \frac{1}{\hat{p}} \frac{z^2_{\alpha/2}}{\varepsilon^2} CV^2_{yU_d} + \frac{(1-\hat{p})}{\hat{p}} \frac{z^2_{\alpha/2}}{\varepsilon^2} \quad (12)$$

Donc

- $$m_d = \frac{1}{\hat{p}} n_d + \frac{z^2_{\alpha/2}}{\varepsilon^2} \frac{(1-\hat{p})}{\hat{p}} \quad (13)$$

Autres enquêtes agricoles ou non

❖ ***Utilisation d'un échantillon complémentaire***

- Il peut être nécessaire de sélectionner un échantillon complémentaire s'il n'est pas possible de renouveler complètement l'échantillon.
- Il est recommandé d'utiliser la base de sondage utilisée pour la sélection de l'échantillon existant et d'utiliser des estimateurs multi-bases *« multiple frames estimators »* appropriés pour les estimations (voir GSARS, 2015b).

Autres enquêtes agricoles ou non

❖ ***Sondage indirect***

- Il y'a des cas où l'unité d'échantillonnage de l'enquête existante est différente de l'exploitation agricole:
  - Enquête agricole aréolaire utilisant des segment ou points
  - Enquête ménage collectant des données agricoles
- Pour de telle situation, il est important de considérer la multiplicité de ces unités avec les exploitations et en tenir compte dans les estimations en utilisant par exemple la méthode généralisée du partage des poids (Lavallée, 2007)

Références

- **Ardilly, P.** (2ème édition actualisée et augmentée, 2006), *Les techniques de sondage*, Editions Technip, Paris**.**
- **Cochran, W.G.** 1977, *Sampling Techniques*. 3rd Edition. John Wiley & Sons: New York, USA.
- **Food and Agriculture Organization of the United Nations (FAO) & United Nations Population Fund (UNFPA)**. 2012. Guidelines for Linking Population and Housing Censuses with Agricultural Censuses: with selected country practices. FAO Publication: Rome.
- **FAO,** 2015, *World Census of Agriculture 2020. Volume 1: Programme, concepts and definitions.* FAO Publication: Rome.
- **FAO**, 2017, SDG Indicator 2.4.1 Proportion of Agricultural Area under Productive and Sustainable Agriculture: methodological concept note. FAO Rome. Available at http://www.fao.org/fileadmin/templates/es/SDG/SDG_2.4.1_Methodological_concept_note.pdf. Accessed on 14/02/2018
- **Fuller, W. A.** 2009, Sampling statistics. Hoboken, NJ: Wiley
- **GSARS,** 2015a, *Guidelines of Integrated Survey Framework*. GSARS Technical Report: Rome.
- **GSARS,** 2015b, *Handbook on Master Sampling Frames for Agricultural Statistics: Frame Development, Sample Design and Estimation*. Global Strategy Handbook: Rome.
- **GSARS,** 2017, *Handbook on the agricultural integrated survey (AGRIS)*. Global Strategy Handbook: Rome.
- **Lavallée, P.** 2007, *Indirect Sampling*. Springer: Ottawa.

IMPROVING AGRICULTURAL & RURAL STATISTICS
Global Strategy

Merci