



# **SESSION 7A**

## **METHODE DE COLLECTE DES DONNEES AGRICOLES**



1

Les directions en charge des statistiques agricoles et l'INS sont tenues de fournir des statistiques agricoles pour répondre aux besoins des utilisateurs.

Quatre sources de collectes sont utilisées pour la collecte:

- la collecte administrative
- les estimations
- l'enquête par sondage
- Le recensement

La collecte administrative et les estimations sont utilisées en cas d'absence d'enquête et de recensement

## 1. COLLECTE ADMINISTRATIVE

- La collecte administrative se fait notamment à l'aide des registres des organisations professionnelles agricoles. Elles aident à constituer une liste des exploitations agricoles et/ou aquacoles. Il est en effet difficile et coûteux de tenir à jour des registres des exploitations.
- La collecte de données administratives est utile surtout pour les cultures encadrées.
  - Pour les cultures encadrées ou organisées en filière, il suffira de contacter toutes les organisations ou les interprofessions de la culture pour avoir les données disponibles entre autres la production; il s'agit très souvent de la production commercialisé, exporté ou transportés.
  - Pour les cultures vivrières, la collecte administrative est difficile.

## ESTIMATION DES DONNEES

4

Deux méthodes d'estimation sont utilisées.

- **Méthode dite de tendance centrale:** pour les séries de données de moins de 15 ans.
- **Modélisation des séries:** elle demande de disposer d'une longue série de données (20 ans);

### a) Estimation par la méthode de tendance centrale

Elle consiste à estimer la production en se basant sur le taux de croissance des 5 dernières années.

- On calcule le taux de croissance selon la formule suivante :

$$\hat{\alpha} = \frac{1}{5} \sum_{t=1}^5 \left( \frac{Y_t}{Y_{t-1}} - 1 \right)$$

- On estime la production de l'année selon la formule suivante

$$Y_t = Y_{t-1} (1 + \hat{\alpha})$$

## b) Estimation de la production par modélisation

5

Elle consiste à rapprocher les variations de la série de données à celles d'un processus d'auto régression AR(p) et on détermine les paramètres (p=nombre de la série).

$$Y_t = \alpha Y_{t-1} + \beta$$

représente un processus AR (P). On détermine les valeurs de la série estimée ; dans notre exemple:

$$\hat{Y}_t = \hat{\alpha} Y_{t-1} + \hat{\beta}$$

En fonction des informations sur le déroulement de la campagne (perspectives de production, pluviométrie, appréciation de la campagne) on corrige l'estimation par :

$$Y_t = \hat{Y}_t + E_n$$

## c) Estimation de la production avec les données d'enquête

6

Culture **pure**

$$P = S_P * R * NB_C$$

- Cultures en **association**

$$P = S_P * R * C * NB_C$$

**P**=production (tonnes), **S<sub>P</sub>**=Superficie en production,

**R**= Rendement (tonne à l'hectare), **NB<sub>C</sub>**= Nombre de Cycles,

**C**= Rapport entre la superficie théorique équivalente en culture pure et la superficie en association

**Estimation de la superficie avec les données d'enquête**

$$S_p = S_T * D_e / D_{ep}$$

*S<sub>p</sub>* la superficie totale en production,

*S<sub>T</sub>* la superficie totale en culture associée,

*D<sub>e</sub>* la densité observée en culture élémentaire et

*D<sub>ep</sub>* la densité en culture élémentaire pure

**NB : Superficie totale  
calculée peut être  
supérieure à la  
superficie physique**

## d) Estimation du rendement avec les données d'enquête

7

Dimension des parcelles élémentaires est fonction de l'homogénéité des cultures sur la parcelle:

- Parcelle est de taille suffisante dès qu'il y a 15 à 25 unités (pieds ou poquets)
  - Culture pure homogène : 3m x 3m (riz) ou 5m x 5m (arachides)
  - Culture pure hétérogène : 10m x 20m (cultures arbustives assez dispersées)
  - Culture en association: 5m x 5m (mil-mais)
- Nombre de carrés fonction de l'homogénéité des cultures sur la parcelle

- Comptage du nombre de pieds
  - Sur toute la parcelle (culture en association non homogène)
  - Sur le carré (culture pure ou en association homogène)
- Evaluer le nombre moyen de pieds à l'hectare (densité observée)
  - Culture pure
  - Chaque culture en association
- Evaluer la superficie totale donnée (sommation des données GPS)
  - Culture pure
  - Culture associée
- Décomposer les superficies par culture

### e) Evaluation du rendement

- **Rendement unitaire moyen** : Production/Superficie (tout le pays)
- **Rendement moyen à l'hectare** : obtenu par prélèvement et pesée
- Effectuer la pesée des récoltes enregistrées sur les carrés de densité/rendement
- Convertir les rendements bruts (problème d'unité)

Cultures	Comptes à la récolte	Pesée Brut Frais	Pesée Brut Sec	Pesée marchande Frais	Pesée marchande Sec
Mil	Epis	Epis		Egrené	Egrené
Sorgho	Epis	Epis		Egrené	Egrené
Maïs	Epis	Epis		Egrené	Egrené
Coton	Capsules	Coton graines	Coton graines		Coton fibre
Niébé				Ecossé	
Arachide		Coques	Coques	Ecossé	Décortiqué
Oignon		Bulbe & tige			
Patate		Tubercules			
Manioc	Bottes	Tubercules			
Pommes de terre	Bottes	Tubercules			
Blé	Bottes	Bottes		Grains	Grains
Riz	Bottes	Bottes	Paddy		Décortiqué

## ENQUÊTE PAR SONDAGE

- les principales questions
  - Quelle est la population cible (les populations ciblées)?
  - Comment constituer l'échantillon ?
  - Quelle taille d'échantillon choisir?
  - Jusqu'où peut on faire confiance aux résultats obtenus sur l'échantillon (**précision des estimateurs**) ?
  - Peut on extrapoler (*généraliser*) les résultats à toute la population ?
- **Les unités : bien les identifier**
  - **L'unité d'échantillonnage** (éléments de la base de sondage): ménages
  - **L'unité d'observation** (unité sur laquelle les données sont recueillies ou dérivées) :
    - ménages : pour les caractéristiques sociodémographiques des membres
    - l'exploitation agricole : activités pratiquées, type, caractéristique

## ▪ **Méthodes de sondage**

La disponibilité de la base de sondage est le critère principal qui détermine la méthode de sondage qui sera adoptée.

**«Dispose-t-on » ou « peut-on constituer une base de sondage? »**

- **Si non : Méthode Empirique**
  - Choix raisonné des enquêtés
  - Pas de calcul de précision
  - Pas de généralisation des résultats (extrapolation)

### **Types de méthode**

- Quotas
- Itinéraires
- Volontariat
- Boule de neige
- Unités types

- **Si oui : Méthode Probabiliste**
  - Sélection aléatoire des enquêtés
  - Estimateurs et précision
  - Généralisation des résultats

### **Types de méthode**

- Sondage aléatoire simple (SAS)
- Sondage stratifié
- Sondage à plusieurs degrés
- Sondage par grappes

## Techniques de sondage empirique (a revoir la page)

- Les méthodes empiriques d'échantillonnage = méthodes de sondages à choix raisonné
- Méthodes pour lesquelles les unités statistiques d'un univers sont choisies afin de rendre l'échantillon représentatif ou respectant un critère particulier lorsqu'il n'y a pas :
  - Pas de base de sondage
  - Pas assez de moyens financiers;
  - Pas de mécanisme de sélection aléatoire des individus de la population
  - Pas d'utilisation des probabilités d'inclusion
  - Pas de possibilité d'extrapoler de façon rigoureuse les résultats à toute la population
  - Pas de calcul d'intervalle de confiance

## a) Méthode des quotas

- L'échantillon est une représentation à l'échelle réduite de l'univers statistique à étudier, du moins pour certains caractères.
- On calcule les différentes proportions des sous populations de l'univers qui sont aussi celles de l'échantillon.
- A chaque enquêteur on donne alors le Nombre d'individus à interroger ainsi que leur répartition par rapport aux critères.

### **Quand l'utiliser ? :**

- On n'a pas de base de sondage
- Pour constituer un petit échantillon
- On a de l'information auxiliaire : les répartitions (quotas)
- On veut construire une image de la population selon les informations auxiliaires
- Faible financement

## b) Méthode des itinéraires

- Elle consiste à indiquer à l'enquêteur, les itinéraires à suivre ainsi que les lieux où il devra s'arrêter pour faire une interview.
- Cette méthode est un bon complément à la méthode des quotas. Les itinéraires sont choisis soit de façon aléatoire, soit comme des « lieux types ».
- Les contraintes d'itinéraires sont alors élaborées de manière à introduire à la fois la représentativité de l'échantillon et l'aléa dans la sélection.

### a) Avantages :

- Permet de prendre en compte les caractéristiques géographiques de l'univers (types de quartiers par exemple) , et donc, d'avoir la meilleure couverture ;
- laisse peu d'initiative aux enquêteurs.
- Son coût est peu élevé et le problème des non-réponses peut être résolu facilement

### b) Inconvénients

- Le problème de non réponse n'est pas très souvent résolu (que faire en cas d'absence ?)
- Le caractère de représentativité dépend
- des "connaissances géographiques" et du discernement de la personne définissant l'itinéraire

## Méthode "boule de neige"

16

- Elle est surtout appliquée pour constituer un échantillon basé sur une caractéristique relativement rare dans l'univers et pour laquelle on dispose de peu d'information quant à leur localisation.
- L'idée est alors de partir d'un individu identifié et de retrouver les autres possédant cette caractéristique à partir des indications que le premier répondant donnera à la demande de l'enquêteur.

### a) Avantages :

- Coûts faibles ;
- Pas besoin de base de sondage ;

### b) Inconvénients :

- Biais de sélection dû aux informateurs initiaux ;
- Difficile à mettre en œuvre ;
- En aucun cas, l'échantillon ne peut être considéré comme représentatif d'une population ;
- Temps de recueil peut être incertain ;
- Difficile d'atteindre la taille voulue ;

## méthode des unités-types

- Elle consiste à subdiviser la population à étudier en un certain nombre de sous-populations relativement homogènes.
  - L'échantillon sera formé d'unités-types provenant de chacune de ces sous populations.
  - nécessite une connaissance préalable de l'univers étudié et pose des problèmes d'extrapolation des résultats
- **a) Avantages :**
    - Très rapide
    - Permet de recueillir des caractéristiques objectives des grandes unités.
    - Produit parfois d'assez bons résultats
  - **b) Inconvénients :**
    - Nécessite une bonne connaissance préalable de l'univers étudié ;
    - Le choix de l'unité type est arbitraire ;
    - Les résultats ne peuvent être généralisés.

**Remarque :** En général, il faut éviter cette méthode

## Observations générales sur la méthode empirique

- **Avantages**

- Peuvent être utilisées quand aucune base de sondage n'est disponible.
- Souvent plus rapides et moins coûteuses que les échantillons aléatoires.

- **Inconvénients**

- La validité des échantillons non aléatoires ne peut être évaluée que subjectivement.
- On ne peut pas calculer des variances d'estimateurs.
- Ils fournissent habituellement des échantillons biaisés.

## 2. Techniques de sondages aléatoires ou probabilistes

- Il y a une base de sondage
- On utilise une sélection aléatoire des individus à enquêter (en général par des logiciels)
- On peut donc déterminer la *précision* des estimateurs
- On peut généraliser

### Type de Techniques de sondages aléatoires ou probabilistes

- Sondage aléatoire simple
- Sondage stratifié
- Sondage à deux degrés
- Sondage par grappes

## Le sondage aléatoire simple (sas)

- **Principe** : Tirer (sélectionner) dans une liste (taille N) les individus à enquêter (taille n)
  - Avec remise (rare)
  - Sans remise (en général)

### Quand :

- Il n'est pas *nécessaire* de *contrôler* l'échantillon par des variables auxiliaires, et surtout
- population est assez **homogène** (du point de vue des informations recherchées).

### Mise en œuvre

- Constituer la base de sondage (liste des individus de la population)
- Définir combien d'individus à tirer
- Définir la méthode de tirage (avec remise ou sans remise)
- Sélectionner le nombre d'unités retenues.

- **Avantage**

- Facile à mettre en œuvre;
- Rapide;
- Pas besoin d'information auxiliaire;
- On peut généraliser;
- On a les intervalles de confiance;
- La répartition et la précision s'améliorent avec la taille « n ».

- **Inconvénients**

- Besoin d'une base de sondage
- Possibilité de forte dispersion (sur le terrain)
- D'où coûts élevés (temps et argent)
- Pas toujours possibles d'obtenir des estimations dans les subdivisions de la population

## Le sondage stratifié

22

- **Principe**
  - Répartir la population en « H » groupes *homogènes* (= **strates**) et
  - sélectionner dans chaque groupe (taille  $N_H$ ) « par un SAS » une liste d'individus à enquêter (taille  $n_h$ )
    - $N = N_1 + N_2 + \dots + N_H$
    - $n = n_1 + n_2 + \dots + n_h$
- **Utiliser quand :**
  - Il n'est pas *nécessaire* de *contrôler* l'échantillon par des variables auxiliaires, et surtout
  - population est **homogène** (du point de vue des informations recherchées)

### Mise en œuvre

- Constituer la base de sondage (liste des individus de la population);
- Les répartir par strate ou groupe;
- Définir combien d'individus tirer;
- Répartir l'échantillon par strate;
- Tirer le nombre d'unités voulu dans chaque strate;

## Avantages

- Amélioration de la précision globale;
- Possibilité d'obtenir des estimations par strate et de comparer les strates;
- Généralisation possible;
- Maîtrise de la dispersion de l'échantillon sur le terrain;

## Inconvénients

- Besoin d'une base de sondage.

## Le sondage à deux degrés

- **Principe :** Tirage en deux étapes :
  - On tire d'abord un échantillon des groupes (**Unités Primaires : UP**), puis dans chacun d'eux
  - On tire un échantillon d'individus (Unités Secondaires US) à enquêter
- **Utiliser quand :**
  - Pas de base de sondage des unités finales (US), mais juste une base d'unités intermédiaires (UP).
  - En supposant que le phénomène est dispersé dans chaque UP, et donc chaque UP dispose de l'information cherchée.
- **Mise en œuvre**
  - Constituer la base de sondage des unités intermédiaires (DR, Villes, Communes, etc. (liste des UP)
  - Déterminer et tirer le nombre d'UP voulu
  - Etablir la liste des unités de chaque UP tirée (Liste des US) *Par un dénombrement*
  - Tirer les US échantillon dans chaque UP tirée
  - NB. Il y a arbitrage entre le nombre d'unités !!!!!!!

## **Avantages**

- Possibilité d'obtenir des estimations par groupe géographique (contenant plusieurs UP) et de faire des comparaisons entre ces groupes
- Généralisation possible
- Meilleure maîtrise de la dispersion de l'échantillon sur le terrain (donc des coûts)
- Pratique et très utilisé dans les grandes enquêtes.

## **Inconvénient**

- Besoin d'une base de sondage
- Besoin d'informations auxiliaires pour constituer les UP
- Moins précis que le SAS
- Coûteux de dresser la liste des US dans chaque UP

## Sondage par grappe

Sondage à deux degrés où : Dans chaque UP on effectue un recensement  
C'est tout !

### RESUME

Méthode	Formation de groupes	Homogénéité des groupes	Tirage des groupes	Tirage US dans les groupes
SAS	Non	Non concerné	Non concerné	Non concerné
Stratifié	Oui	Oui	Non	Oui
Deux degrés	Oui	Pas besoin	Oui	Oui
Grappes	Oui	Non, surtout pas	Oui	Non