

Échantillonnage ou méthodes de sondage

R.TALHI
Service de Biostatistique
Faculté de Médecine d'Oran

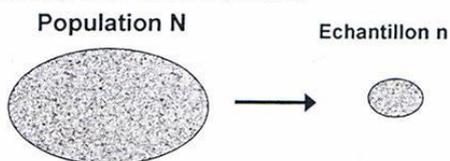
Introduction

Les enquêtes épidémiologiques exhaustives touchent toutes les unités d'une population : ex. le recensement d'une pop. (trop longues et trop coûteuses).

Les enquêtes épidémiologiques qui procèdent par une méthode de sondage, utilisent un sous ensemble, un **échantillon** (*Eng: sample*).

Définition

L'échantillonnage ou le sondage est l'ensemble des techniques qui consistent à sélectionner l'échantillon qui représente au mieux la population source et dont les caractéristiques lui sont très proches.



«L'échantillonnage, c'est regarder attentivement une partie d'une population afin d'en apprendre plus sur elle»

Concepts de base

Population mère ou population source : ensemble de la population à partir de laquelle les sujets de l'étude proviennent.

Population cible ou population d'étude : population visée par l'étude.

Échantillon : sous ensemble de la population cible. un échantillon doit être l'image réduite mais fidèle de cette population

Un échantillon tiré au sort (aléatoire) est **statistiquement représentatif**

Les méthodes d'échantillonnage

o Il existe deux principales méthodes :

- Les méthodes empiriques non probabiliste
- Les méthodes aléatoires ou probabiliste.

5

I- Les méthodes empiriques

On les utilise lorsqu'on ne dispose pas de base de sondage. Ils sont rapides à effectuer et ne coûtent pas cher. En revanche ils sortent du cadre probabiliste et ne permettent pas de calculer la précision des estimateurs. Ils sont particulièrement utilisés dans les sondages d'opinion, marketing et par les instituts de sondage. Ces techniques sont peu utilisées en médecine.

Des biais peuvent apparaître, toutes les unités n'ont pas la même probabilité d'être choisies.

L'échantillon ainsi constitué **n'est plus représentatif.**

6

Principales méthodes empiriques :

- 1- La méthode des itinéraires
- 2- La méthode des quotas
- 3- La méthode des Unités-types
- 4- Volontariat

1-La méthode des itinéraires

Principe :

Itinéraire pré-établi : enquête dans les maisons situées sur la gauche ou à droite de l'enquêteur selon un parcours pré-établi.

Méthode

Un point de départ pris au hasard

L'enquêteur suit l'itinéraire pré-établi

Intérêt

Technique adaptée dans les régions et les localités ne disposant pas de fichier des populations (zones éparses)

2-La méthode des quotas

Principe :

On élabore un échantillon dans lequel sont choisies des quotas où sont reproduits plus ou moins fidèlement, les mêmes variables (âge, sexe, catégorie socio- professionnelle, le lieu de résidence ...) que ceux de la population générale.

On pose l'hypothèse que les paramètres mesurés sur cet échantillon seront identiques dans la population

Inconvénient :

L'enquêteur est laissé libre dans le choix des individus, sa seule contrainte est de respecter les quotas

Risque de sur-représentation des personnes facilement accessibles.

II. Les méthodes aléatoires

Les unités ou sujets qui constituent l'échantillon à étudier sont tirées au sort.

Tous les individus de la population ont ainsi la même probabilité d'appartenir à l'échantillon.

Cela permet d'utiliser les lois de probabilité dans l'exploitation des données et de généraliser les résultats obtenus à l'ensemble de la population.

II. Les méthodes aléatoires

- 1-Sondage élémentaire
- 2-Sondage systématique
- 3-Sondage à plusieurs degrés
- 4-Sondage stratifié
- 5-Sondage en grappes

1-Sondage élémentaire

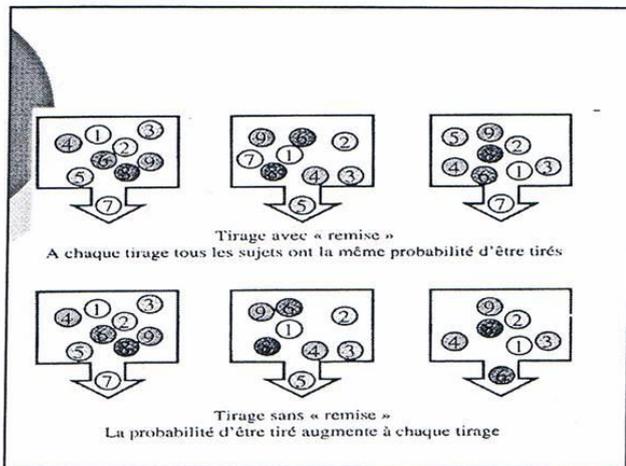
Principe :

C'est le principe de la loterie, qui consiste à prendre au hasard des unités numérotées au préalable dans la population d'étude

Base de sondage : liste numérotée

Après avoir fixé la taille n de l'échantillon, on tire au sort les numéros des individus qui constitueront l'échantillon. Chaque individu a la même probabilité d'être sélectionné

Tirage avec remise – tirage sans remise



En pratique
 tirage au sort : à la main, table de nombre au hasard, logiciel informatique.

Exemple: Si l'on désire tirer au sort 100 individus dans une population de 2000 unités, on attribue un numéros de 1 à 2000 aux sujets constituant la population. A partir de la table des nombres au hasard, on prend tous les chiffres successifs pour atteindre le nombre d'unités de l'échantillon, en utilisant un cheminement horizontal, vertical sur la table.

Table des nombres au hasard

26099	65801	69870	84446	58248	21282	56938	54729	67757	71874	61692	80001
21430	02305	59741	34262	15157	27545	08774	29689	42245	51903	69179	96682
91819	60812	47631	37294	92028	56850	83380	05912	29830	37612	15593	73198
33912	37996	78967	57201	66916	73998	54289	07147	84313	63610	61475	26980
23804	54972	72068	19403	53756	04281	01570	41701	30382	54647	06077	29354
95704	75928	21811	24159	77787	38973	82178	46802	90245	01805	23906	96559
92834	52941	88301	22127	23459	40229	74678	21859	98645	16178	60063	59284
16279	48003	44634	08623	32752	40472	81808	32980	80660	98391	62243	19678
39551	18398	36918	28628	82072	04854	52809	86608	68017	11120	28638	72850
62249	65757	12273	91261	96983-	15082	83851	77682	81728	84541	99891	01585
96711	29712	02877	70955	59693	26838	89052	39061	99811	69831	47234	93263
47386	17462	18874	13407	62899	78937	90525	25033	56358	78902	47008	72488
50230	63237	94083	93634	71652	02656	57532	60307	91619	84980	62458	09703
78397	66179	46982	67619	39254	90763	22116	33646	17545	31321	65772	86506
09811	82848	92211	68645	15068	56898	87021	40115	27524	42221	88293	67592

2-Sondage systématique

Principe :

- il faut que les individus soient classés suivant un ordre quelconque
- ne nécessite pas la numérotation des individus de la population
- on détermine un **pas de sondage** qui est le rapport entre la taille de la population N et la taille désirée de l'échantillon n.

Le pas est donc égal à N/n
 tirer au sort le premier individu entre 1 et N/n
 puis sélection des suivants par intervalles fixes

Exemple : pour prélever un échantillon sur une population de 2000 unités avec un pas de sondage de 10, on prend un nombre aléatoire compris entre 1 et 10 : 5, puis l'échantillon commencera par (5), (5+10), (15+10), (35), ...

3-Sondage à plusieurs degrés

Principe

partition de la population en groupes (unités primaires) puis éventuellement en sous groupes d'unités secondaires puis tirage au sort élémentaire ou systématique

On peut obtenir un échantillon en utilisant plusieurs tirages au sort successifs .

Exemple:

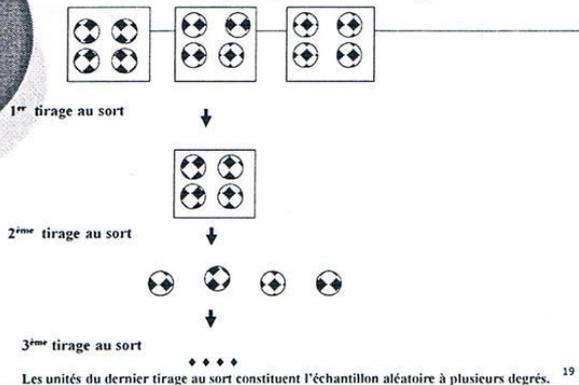
dans un sondage à trois degrés, on effectue un premier tirage au sort des communes (unités primaires) à l'intérieur desquelles on tire un échantillon de districts, le troisième tirage concernera les ménages.

Sur le plan pratique,

ce type de sondage permet de diminuer les déplacements des enquêteurs.

18

Sondage à plusieurs degrés



19

4- Sondage stratifié

La stratification consiste à subdiviser la population à étudier en groupe homogènes (appelés strates), selon les modalités d'un caractère.

La population à étudier est donc préalablement classée en différentes « strates ».

Les strates sont définies par rapport à des caractères de la population (secteur d'habitat, classes économiques, tranches d'âge, par sexe ...).

Principe

- découpage de la population en strate
- dans chaque strate effectuer un sondage aléatoire

20

La stratification constitue un moyen pour accroître la précision du sondage. Le gain de précision dû à la stratification est d'autant plus grand qu'il existe une liaison plus forte entre le critère servant à la stratification (l'âge par ex.) et la variable étudiée.

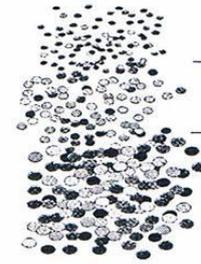
Il existe **deux types** de stratification :

- La stratification à priori
- La stratification à posteriori

Sondage élémentaire



Population



Sondage stratifié



5-Le sondage en grappes

Grappe :

groupe d'unités de la population répartie en lieu (maison, groupe de maison, service hospitalier)

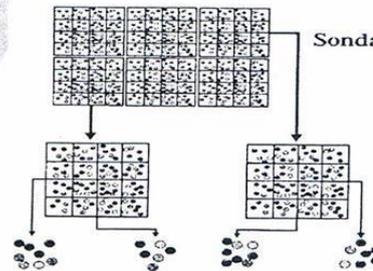
Principe:

On effectue un tirage au sort des différentes grappes qui constituent la population, puis on mène l'enquête auprès de tous les individus composant les grappes.



23

Sondage en grappes



grappes

Comment déterminer la taille d'un échantillon ?

- Cas enquête descriptive : nécessité de
 - connaître une approximation du caractère étudié
 - définir un degré de précision
- Calcul du nombre de sujets nécessaires:

$$n = \frac{\varepsilon^2 p(1-p)}{i^2}$$

ε : écart-réduit (1,96 pour un risque $\alpha = 5\%$)
 p : prévalence estimée
 i : précision désirée

25

Exemple:

On veut évaluer la fréquence d'une maladie dans une population . Cette fréquence est de l'ordre de 10%

La précision désirée est de 2% pour un risque consenti de 5%.

$$n = [(1,96)^2(0,1)(1- 0,1)] / (0,02)^2 = 900$$

Le nombre de sujets nécessaires dans notre cas est de **900** individus.

Conclusion

Le choix d'un échantillon représentatif est une tâche difficile, mais l'on doit se rapprocher au maximum, en utilisant la méthode d'échantillonnage la mieux appropriée, la plus pratique, et la moins coûteuse.

Il est essentiel de veiller aux erreurs et aux biais aux différents moments.

Il est nécessaire également de veiller au respect de la **représentativité de échantillon** au cours du temps, et tout au long de l'enquête (sujets perdus de vue...).

27