



COURS D'ÉPIDÉMIOLOGIE ORGANISÉS PAR L'OMS

Module d'auto-apprentissage - cours en français

(préparé par M. T. K. Sundaresan)

Sujet général : ECHANTILLONNAGE

Sujet particulier : TECHNIQUES DE SONDAGE



Rapport avec les fonctions de l'épidémiologiste

Les enquêtes par sondage constituent un moyen efficace et économique d'obtenir des informations sur certaines questions importantes en épidémiologie, par exemple sur la prévalence de maladies déterminées. Il existe plusieurs techniques de sondage entre lesquelles on peut choisir en fonction des situations pratiques et des objectifs spécifiques de la recherche entreprise. Dans une situation donnée, l'une d'entre elles pourra être plus efficace et plus économique qu'une autre. Il est donc très utile de connaître ces techniques, leurs avantages et leurs inconvénients, pour choisir celle qui correspondra le mieux à la situation.

Objectif éducationnel (Réf. N° 5.6)

Après avoir étudié le présent document et fait les exercices qui le terminent, vous devez pouvoir choisir, pour la plupart des enquêtes épidémiologiques, la technique de sondage appropriée.

Conditions requises pour comprendre le sujet

1. Comprendre la notion de sondage aléatoire (objectif éduc. N° 5.2).
2. Comprendre les termes "précision" et "exactitude" (objectif éduc. N° 5.3).
3. Comprendre la notion d'"intervalle de confiance" (objectif éduc. N° 5.4).

1. Préambule

L'épidémiologiste désireux de procéder à une enquête par sondage doit souvent faire un choix entre plusieurs techniques de sondage. Parmi les considérations qui interviennent dans le choix de la technique la plus appropriée, il convient de mentionner les suivantes : existence d'une base de sondage appropriée, facilité des opérations sur le terrain, précision acceptable des estimations, subdivisions de la population pour lesquelles des estimations distinctes seront nécessaires, etc. Il existe de nombreuses techniques de sondage, mais seules les suivantes seront examinées ici :

The issue of this document does not constitute formal publication. It should not be reviewed, abstracted or quoted without the agreement of the World Health Organization. Authors alone are responsible for views expressed in signed articles.

Ce document ne constitue pas une publication. Il ne doit faire l'objet d'aucun compte rendu ou résumé ni d'aucune citation sans l'autorisation de l'Organisation Mondiale de la Santé. Les opinions exprimées dans les articles signés n'engagent que leurs auteurs.

1. sondage aléatoire simple;
2. sondage stratifié;
3. sondage à deux degrés; et
4. sondage par grappes.

Sondage aléatoire simple

Définition : On parle de sondage aléatoire simple lorsqu'un échantillon de taille n est tiré d'une population de taille N de telle manière que tous les échantillons possibles de taille n aient la même chance d'être choisis.

Exemple On veut déterminer la prévalence d'une maladie infectieuse parmi les 1000 habitants d'un village ($N = 1000$). Compte tenu du degré d'exactitude souhaité pour l'estimation, il sera nécessaire de rechercher la présence de la maladie chez 100 habitants ($n = 100$). On dresse une liste de tous les habitants (base de sondage). Puis, à l'aide d'une table de nombres au hasard, on constitue un échantillon de 100 personnes de telle manière que chacun des habitants (c'est-à-dire chacune des unités de la liste) ait une chance sur 10 (n/N) d'être choisi.

Avantages

1. Simplicité du plan d'échantillonnage.
2. Il n'est pas nécessaire de connaître la distribution de la variable étudiée au sein de la population.

Inconvénients

1. Exige une base de sondage qui peut être coûteuse ou même impossible à obtenir.
2. Les frais de déplacement peuvent être élevés si la population est dispersée sur un vaste territoire.

Sondage aléatoire stratifié

Définition : Dans le sondage aléatoire stratifié, la population est d'abord divisée en groupes distincts appelés strates; on effectue ensuite dans chaque strate un échantillonnage aléatoire simple.

Exemple Reprenons l'exemple donné plus haut au sujet du sondage aléatoire simple : avant de tirer l'échantillon de 100 personnes, on divise toute la population en groupes d'âge (stratification par âge). On tire ensuite un échantillon aléatoire simple dans chaque groupe d'âge (strate). Le nombre total d'examen sera encore de 100, mais la prévalence de la maladie pourra être étudiée par groupes d'âge et la précision de l'estimation globale de la prévalence sera donc améliorée. Cet accroissement de la précision vient de ce que l'échantillonnage aura été fait au sein de groupes plus homogènes (si l'on suppose qu'il existe une corrélation entre l'âge et la prévalence).

Avantages

1. Il est possible d'obtenir une estimation pour chaque strate de la population, et non pas seulement une estimation globale comme dans le cas du sondage aléatoire simple.
2. Il en résulte un gain de précision de l'estimation globale de la population à condition que la méthode de stratification ait été correctement choisie.

Inconvénients

1. Il faut disposer d'une base de sondage.
2. Il faut obtenir certaines informations préliminaires sur la ou les variable(s) à utiliser pour la stratification, ce qui peut être coûteux ou même irréalisable.
3. Si la stratification pour certaines variables permet d'améliorer la précision de l'estimation concernant un paramètre (la prévalence d'une maladie dans l'exemple ci-dessus), elle risque de donner en revanche des estimations assez médiocres pour les autres caractères qui doivent être étudiés en même temps (par exemple, la fécondité).

Sondage par grappes

Définition : Un sondage par grappes est un sondage aléatoire simple dans lequel chaque unité d'échantillonnage est un ensemble ou groupe d'éléments appelé grappe.

Exemple Poursuivons avec l'exemple ci-dessus. Si l'on ne dispose pas d'une base de population, on peut réaliser le sondage comme suit : dresser une liste des unités d'habitation, beaucoup plus facile à obtenir qu'une liste complète des individus. Tirer ensuite un échantillon d'unités d'habitation par sondage aléatoire simple et examiner tous les individus de chaque unité tirée.

Avantages

1. Réduction des frais de déplacement.
2. Une base complète de population n'est pas nécessaire.

Inconvénients

1. Il faut tenir compte de la variabilité au sein de chaque grappe et entre les grappes. Si la variabilité au sein d'une grappe est faible, on risque d'obtenir une estimation d'une précision inutile au sein d'une même grappe avec le même nombre d'examen et par contre, une précision insuffisante pour l'estimation globale.

Sondage à deux degrés ou sous-échantillonnage

Définition : Dans le sondage à deux degrés, on tire d'abord un échantillon de groupes, puis un sous-échantillon d'éléments dans chacun des groupes choisis.

Exemple Conservons toujours l'exemple ci-dessus. Les enquêteurs décident d'utiliser la liste des unités d'habitation, parce qu'ils ne disposent pas de la base de population, mais ils souhaitent couvrir davantage d'unités d'habitation tout en conservant le même nombre d'examen. On commence donc par tirer un échantillon aléatoire simple d'unités d'habitation puis un échantillon aléatoire simple des individus vivant dans chacune des unités d'habitation choisies.

Avantages

1. Couverture plus large. Comme on suppose que les personnes vivant dans une même unité d'habitation sont semblables en ce qui concerne la quantité mesurée, il n'est guère utile de mesurer la même chose tant de fois.
2. Une base complète de population n'est pas nécessaire.
3. Réduction des frais de déplacement.

Inconvénients

1. L'inclusion dans l'échantillon d'une partie seulement des membres d'un ménage ou d'un groupe peut parfois donner lieu à des difficultés.
2. Le caractère incomplet des informations sur la distribution des quantités mesurées aux différents stades peut aboutir à des résultats médiocres.

EXERCICES D'AUTO-CONTROLE*

A. Il est prévu de faire une enquête pour déterminer les taux de fécondité dans un vaste territoire pour lequel il n'existe pas de listes exactes des habitants; par contre, on possède une bonne documentation sur les villages et les villes de ce territoire.

1. Proposez deux plans de sondage possibles.
2. Indiquez deux avantages que présente chacun par rapport à l'autre.

B. On désire connaître le taux de morbidité dans la population d'une ville. On dispose des résultats d'un recensement récent, notamment en ce qui concerne l'âge et le sexe des habitants.

1. Proposez deux plans de sondage possibles.
2. Indiquez deux avantages que présente chacun par rapport à l'autre.

C. On désire estimer la prévalence de la tuberculose dans un pays. Le dernier recensement remonte à 10 ans au moins et l'on sait que certains mouvements de population ont sans doute eu lieu depuis. On pense également que la prévalence de la maladie n'est probablement pas la même dans les zones urbaines et dans les zones rurales. En fonction de la situation dans votre propre pays, formulez des hypothèses sur les facteurs opérationnels et proposez un plan de sondage.

* Les réponses figurent à la page suivante.

REPONSES

A.1. Sondage par grappes et sondage à deux degrés

2. Avantages du sondage par grappes par rapport au sondage à deux degrés :

- diminution des frais de déplacement (moins de villages à visiter);
- plan de sondage plus simple;
- exige une moindre connaissance de la distribution des variables à mesurer;

Avantages du sondage à deux degrés par rapport au sondage par grappes :

- couverture plus large (permet d'obtenir des estimations pour un plus grand nombre de villes et de villages);
- permet plus de souplesse car on peut modifier les fractions sondées pour réaliser des économies si l'on a certaines connaissances préliminaires sur la variabilité entre villages;

B.1. Sondage aléatoire simple et sondage stratifié

2. Avantages du sondage simple

- plan de sondage plus simple;
- il n'est pas nécessaire d'avoir une connaissance préalable de la distribution des variables;

Avantages du sondage stratifié

- il permet de fournir des estimations par âge et/ou sexe;
- il permet d'obtenir des estimations globales plus précises;

C. Supposez qu'il soit plus efficace, sur le plan opérationnel,

1. d'examiner tous les membres du ménage une fois qu'un ménage a été choisi;
2. d'immobiliser le camion et l'équipe mobile de radiologie dans un secteur pendant au moins une semaine;
3. supposez en outre : qu'on peut examiner 50 personnes en une journée, et
4. qu'il faut examiner au moins 20 000 personnes pour obtenir une estimation d'une précision acceptable.

On peut recommander, entre autres, la procédure suivante.

Etant donné les facteurs opérationnels ci-dessus, il serait préférable d'adopter un sondage par grappes à plusieurs degrés, stratifié en zones rurales et zones urbaines. Supposez qu'on décide d'examiner 10 000 personnes en milieu rural et 10 000 personnes en milieu urbain. Si l'on prend comme exemple les zones rurales, les 10 000 personnes pourraient être considérées comme la population de 40 grappes composées chacune de 250 individus (50 personnes par jour pendant cinq jours de la semaine). En supposant que la dimension moyenne des ménages soit de cinq personnes, cela représenterait 50 ménages dans chaque grappe.

Compte tenu de l'ancien recensement, le pays peut être divisé en fonction de ses grandes divisions administratives et les 40 grappes peuvent être allouées à ces divisions proportionnellement à la population de ces dernières. Au sein de chaque grande division, on pourra choisir au hasard (deuxième degré) des divisions plus petites dans lesquelles les grappes devront être

prises. Si les divisions sont grandes, peut-être faudra-t-il prévoir des degrés supplémentaires.

Une fois qu'aura été atteint le dernier degré, par exemple une subdivision peuplée d'environ 1000 habitants, on procédera à un recensement complet de cette subdivision et l'on choisira un ménage au hasard. Cinquante ménages consécutifs choisis selon un ordre déterminé à l'avance donneront la grappe requise.

* * *