

Cours n° 13

Intitulé du cours	L'échantillonnage
Points à développer	1. Définitions 2. Types d'échantillon 3. La représentativité

1. Définitions

L'échantillonnage est une technique statistique qui consiste à sélectionner un petit échantillon au sein d'une population totale pour obtenir des résultats mesurables pouvant être extrapolés à l'ensemble de la population, c'est-à-dire choisir un échantillon aléatoire qui est représentatif de l'ensemble du groupe. En sociologie comme il est difficile voire même impossible dans certains cas d'interroger l'ensemble de la population mère, il est souvent nécessaire de procéder à l'échantillonnage (la façon avec laquelle on choisit un échantillon).

L'échantillon est une partie d'un tout, ou opération qui consiste à choisir quelques éléments d'un tout. La notion d'échantillon est relative, chaque discipline à sa façon de procéder à l'échantillonnage. L'échantillon est constitué par tirage au sort à partir d'une base de données qui contient une liste précise de la population de référence. Pour tirer un échantillon, il faut qu'il y ait une population d'étude, qui représente l'ensemble des éléments qui ont des caractéristiques communes et sur lesquels le chercheur projette pour choisir un groupe à étudier. Elle est souvent appelée « population mère ».

2. Types d'échantillon

Les sociologues ont souvent tendance à présenter l'échantillonnage sous deux formes principales: les échantillons probabilistes et les échantillons non probabilistes.

A/ L'échantillon probabiliste

Il consiste à donner à tous les éléments de la population mère la même probabilité et la même chance d'être sélectionnés. De ce fait, il faut avoir des liste exhaustive est nominative de toute la population mère. Les résultats obtenus à travers ce type d'échantillon seront inférer à l'ensemble de la population mère. Il y a plusieurs sortes d'échantillons probabilistes:

1- Echantillon aléatoire simple : ce type d'échantillon exige une liste de tous les éléments de la population mère numérotée de 01 à N, le tirage se fait au hasard, l'idéal est de mettre tous les éléments de la population mère dans le même socle, puis passer au tirage au hasard comme le jeu d'une loterie (tous les éléments ont la même chance d'être sélectionnés).

2- Echantillon au hasard stratifié : dans ce cas, la population mère est divisée selon certains critères comme l'âge, le sexe, la religion...etc. Les strates sont des catégories construite selon

un critère particulier de la recherche, le tirage se fait au hasard dans chaque strate à condition de respecter les proportions de chaque strate dans la population mère.

3- Échantillon systématique : appeler aussi par intervalle, ce qui signifie qu'il y a un intervalle à respecter entre chaque unité et l'autre. La procédure est de numéroter les unités de la population mère de 01 à N, puis on détermine l'intervalle d'échantillonnage (K), en divisant le nombre de la population mère sur la taille de l'échantillon. Ex : si la population mère est constituée de 5000 personnes, l'échantillon souhaité est de 100 personnes, donc l'intervalle est 50 ($k=5000/100=50$).

4- Échantillon en grappe : on utilise cette technique quand il n'y a pas une liste de tous les éléments de la population mère ou quand on est contraint par des considérations de temps et de moyens. Au lieu de projeter sur les individus on projette sur des groupes les contenant.

B/ L'échantillon non probabiliste

On choisit souvent ce type d'échantillon dans les cas suivants :

- Quand il n'y a pas une base de données, c'est-à-dire on ne possède pas au préalable des listes devant contenir tous les éléments de la population mère.
- Lorsqu'on veut approfondir la compréhension de certains comportements sans prendre en considération la représentativité.
- Dans l'étude de cas.
- Lorsqu'il est difficile pour le chercheur d'avoir un échantillon probabiliste faute de moyens et de temps.

Il y a plusieurs sortes d'échantillons probabilistes:

1. l'échantillon accidentel : On utilise ce genre d'échantillon quand il n'y a aucune base de données sur la population mère, donc on peut interroger les premiers éléments sur lesquels on tombe accidentellement (par hasard).

2. L'échantillon typique : appeler aussi *échantillons exemplaire*, dans ce cas le chercheur constitue un échantillon d'individus considérés comme des modèles qu'ils vont représenter la population étudiée.

3. l'échantillon par quotas : le chercheur doit avoir au préalable une idée sur les caractéristiques de la population mère de façon à reproduire les mêmes caractéristiques dans son échantillon, en plus il doit prendre un cotas proportionnelle par rapport à la population mère.

4. Échantillon boule de neige : on utilise ce type d'échantillon quand la population mère est inconnue, donc l'accent sera mis sur une personne ou plusieurs personnes (informateurs) pour nous mettre en contact avec d'autres. Dans ce cas de figure le seul moyen est d'avoir des connaissances pour arriver à constituer un échantillon.

5. échantillon volontaire: on demande à un ensemble de personnes de faire partie volontairement d'un échantillon d'une étude avisée à l'avance. Pour plus d'objectivité le chercheur doit passer une annonce à travers les moyens de communication possibles pour constituer son échantillon.

3. La représentativité

La représentativité d'un échantillon, c'est sa capacité à offrir les mêmes caractéristiques que celles de la population qu'il est supposé représenter. Au moins sur les critères jugés structurants par rapport à ce que l'on cherche à éclairer. Par exemple, pour la population donnée, les critères de représentativité seront généralement : le sexe, l'âge, la CSP, la région, le niveau d'instruction...etc.

Dans un premier sens la représentativité d'un échantillon qui soit aléatoire ou raisonné dépend de sa taille, de la fiabilité de la précision, de la base de sondage et de la rigueur du protocole de recueil des informations. Mais c'est la représentativité statistiques des sondages qui évoquent surtout le mot « représentativité » dans la mesure se fonde sur la loi des grands nombres et le calcul des probabilités et suppose donc une procédure de sondage aléatoire.