


METHODOLOGIE DE L'ENQUETE

■ **Enquête systématique et enquête partielle**

Dans le cadre d'une enquête dite systématique, on a besoin de mobiliser un grand nombre d'enquêteurs durant une longue période. Dans ce cas, le dépouillement des réponses ainsi que l'exploitation des résultats risquent de devenir des opérations très lourdes. L'enquête systématique ne peut être recommandée que lorsque le volume de la population prise en considération est réduit.



Dans **le** but de contourner les difficultés d'ordre économique, on préfère interroger uniquement une partie de la population et l'on effectue ainsi, ce qu'on appelle une **enquête partielle** ou **un sondage**.

La partie extraite de la population totale s'appelle **échantillon** et la population totale s'appelle **univers**. Pour garantir l'exactitude des résultats du sondage et s'assurer de leur validité, le chercheur est amené à extraire un **échantillon représentatif**.


■ La représentativité de l'échantillon

- Un échantillon **représentatif** peut être considéré comme une sorte de modèle réduit de l'univers.
- Si le principe de la représentativité est respecté, on peut, à quelques erreurs près, "**erreurs d'échantillonnage**", parvenir aux mêmes résultats en travaillant sur un échantillon réduit qu'en travaillant sur l'univers total.
- Il est vrai qu'un **échantillon représentatif à dix pour cent** simplifie la perception de la réalité qui est souvent fort complexe tout en permettant l'apparition des **constantes** ou des **tendances de base**. En effet, dans un petit échantillon, les problèmes de base qui concernent tous les éléments de l'univers ou du moins la majorité, apparaissent nécessairement, tandis que les problèmes secondaires, considéré généralement comme des **cas statistiques aberrants**, tendent à disparaître.



Deux questions fondamentales se posent :

- - Comment doit-on extraire les éléments de l'échantillon pour que celui-ci soit suffisamment représentatif ?
- - Comment peut-on estimer l'ampleur des erreurs de l'échantillonnage pour s'assurer quant à la validité des conclusions et pour justifier la généralisation des résultats ?.



La réponse à la deuxième question concerne les **seuils de confiance** ou **seuils de probabilité**, seuils qui précisent les probabilités pour qu'un **événement aléatoire** quelconque apparaisse ou n'apparaisse pas ou pour qu'une conclusion donnée soit plus proche de la fausseté ou de la vérité.

En ce qui concerne la première question nous devons préciser qu'un échantillon est dit **représentatif** lorsqu'il possède les mêmes caractéristiques de l'univers dans les mêmes proportions que lui.



C'est dans ce sens qu'un échantillon représentatif peut être assimilé à un modèle réduit de l'univers.

Les variables comme l'âge, le sexe, la religion, l'état civil, le niveau socioéconomique ou catégorie socioprofessionnelle (CSP), le niveau intellectuel, le grade, la région géographique etc., peuvent constituer les caractéristiques de l'univers.

Ces variables déterminent les réponses des individus aux questions et par conséquent, nous devons en tenir compte au moment de l'extraction de l'échantillon et faire de sorte qu'elles y apparaissent dans les mêmes fréquences que celles qui caractérisent l'univers lui-même.



La constitution d'un échantillon représentatif doit obéir à des principes probabilistes simples : ce sont les lois mathématiques du hasard qui régissent la représentativité.

La règle peut être formulée de la façon suivante :

l'extraction d'un échantillon doit être faite de sorte que chaque élément de l'univers ait la même chance d'apparaître que chacun des autres éléments. Une extraction qui ne se conforme pas à ce principe probabiliste aboutit à un échantillon non représentatif qu'on appelle aussi un "**échantillon préjugé**" ou encore un "**échantillon biais**".

■ Les techniques de l'échantillonnage

■ L'échantillonnage au hasard simple ou probabiliste

Le mot hasard est utilisé dans le sens de contrôle des opérations au moment de l'extraction des éléments de l'échantillon.


L'échantillonnage au hasard simple est une technique qui s'impose lorsque le chercheur ne dispose que d'un répertoire nominatif de tous les individus qui composent l'univers. Le chercheur ignore totalement la description de l'univers en termes de strates ou de catégories telles que l'âge et le sexe par exemple. A cet effet, il ne dispose pas de statistiques officielles. Il connaît seulement les noms et les adresses des sujets.



■ Il existe d'autres possibilités plus pratiques :

-La première consiste à utiliser une table de nombres au hasard. On donne à toutes les unités de l'univers des numéros allant de 1 à N, mais au lieu d'utiliser une urne pour effectuer le tirage au sort, on repère dans la table un numéro au hasard, puis on continue à lire les autres numéros en respectant soit le **sens horizontal**, soit le **sens vertical**, soit enfin le **sens oblique** et en partant de **gauche à droite** ou **inversement**.

-Ce sont ces nombres lus dans la table qui vont être retenus en tant qu'éléments de l'échantillon pourvu qu'ils ne dépassent pas le volume ou la taille de l'univers.



En cas de dépassement on doit tout simplement ignorer le nombre apparu au lieu de l'enregistrer. **Comme les nombres dans la table sont des nombres exposés au hasard, le tirage au sort effectué de cette façon représente un échantillonnage probabiliste.**


-La seconde solution consiste à confier toute l'opération à un ordinateur.

-Le **troisième procédé** est connu sous le nom **d'échantillonnage par intervalles réguliers.**

L'échantillonnage *par* intervalles réguliers

C'est un tirage probabiliste qui obéit au même principe que celui de l'échantillonnage au hasard simple. La seule différence réside dans la simplification des opérations manuelles.

Une fois que les unités sont numérotées, **le procédé consiste à choisir les matricules par intervalles réguliers**. On commence par considérer les numéros suivant leur **ordre croissant** allant de 1 à N (ou décroissant de N à 1) et, ensuite, on relève les matricules qui se succèdent à des **distances égales**.



Par exemple, on peut retenir suivant l'ordre croissant le **cinquième sujet** (dont le matricule est 5), ensuite le **dixième**, le quinzième, le vingtième et ainsi de suite jusqu'à ce que l'échantillon soit constitué.


Le **choix du premier matricule peut varier** puisqu'on peut choisir, à la place des sujets précédents, les sujets numérotés consécutivement 4, 9, 14, 19 etc. ou encore les numéros 3, 8, 13, 18 etc. Mais les **distances demeurent constantes** : dans les exemples précédents, la **raison arithmétique** est toujours égale à 5 et la suite est croissante.

■ **La raison** varie quand on passe d'une enquête à une autre, en fonction des dimensions de l'univers et de l'échantillon que l'on cherche à extraire.

■ **Exemple:**

on peut imaginer qu'il s'agit d'un univers de **2000 sujets** et d'un échantillon qui doit contenir **400 sujets (20% ou 1 sujet pour 5)**. La suite arithmétique qui commence à partir de 5 ou de 4 ou encore de 3 et dont la raison vaut 5 remplit les conditions : **seul le premier matricule est désigné au hasard (entre 1 et 5)**; les autres sont obtenus en ajoutant chaque fois la raison 5.

On obtient cette raison en divisant 100 par 20, ce qui correspond au pourcentage 20%.



■ La méthode au hasard simple et la méthode par intervalles réguliers ne sont applicables que lorsque le chercheur ignore tout de la composition de l'univers. Lorsque cette dernière est connue, l'échantillonnage par stratification est plus justifiable que les deux méthodes précédentes.

L'échantillonnage par stratification

Contrairement à la méthode au hasard simple, la **stratification** suppose, en plus des enseignements élémentaires, que la composition de l'univers en termes de **variables indépendantes** (sexe, âge, état civil, catégorie socioprofessionnelle etc.) est bien connue. Il est possible alors de constituer un échantillon représentatif en tirant isolément des **sous-échantillons** dont les dimensions seraient proportionnelles à celles des strates significatives à l'intérieur de l'univers.

La dénomination **strate** ne doit être confondue ni avec celle de **catégorie** ni avec celle de **variable**.

Variable: tout élément qui dépend de certaines **circonstances** et qui prend, selon ces circonstances, soit des **valeurs quantitatives** différentes, soit des **modalités qualitatives** diverses.



Catégorie: une modalité parmi les modalités à partir desquelles une variable qualitative se compose.

Par exemple, le **sexe est une variable qualitative** qui comprend deux catégories, celle des hommes et celles des femmes. De même, la variable niveau socio-économique peut être considérée comme constituée de trois catégories, celle des gens aisés, celle de la classe moyenne et celle des gens de niveau plus bas.

La **notion de strate** signifie **une subdivision de l'univers en fonction de la combinaison ou de l'intersection d'un certain nombre de variables indépendantes** qui sont jugées importantes par le chercheur.

Par exemple, le **sous-ensemble des femmes âgées et riches** constitue une strate définie par trois variables indépendantes, à savoir le **sexe, l'âge et le niveau socio-économique**.

Le nombre de **strates** est alors égal au produit des nombres de catégories appartenant aux variables retenues.

Dans l'exemple précédent, en supposant que chacune des trois variables se compose de deux catégories (hommes et femmes en ce qui concerne le genre, jeunes et adultes en ce qui concerne l'âge et, enfin, riches et pauvres pour ce qui est du niveau socio-économique), le nombre de strates est égal à 8 qui est le produit des six catégories (**$2 * 2 * 2 = 8$**).

Technique de stratification

- Commencer par découper l'univers en un certain nombre de dimensions que l'on juge importantes et significatives telles que l'âge, le sexe et la profession des sujets,

- Ensuite, on tire au sort en appliquant la méthode au hasard simple ou par intervalles réguliers à l'intérieur de chaque strate définie au préalable par le chercheur.

- Le principal objectif visé par l'intermédiaire de la stratification réside dans la possibilité d'obtenir un échantillon le plus représentatif et le plus fidèle qui soit. Seulement, la stratification n'est possible que si la composition de l'univers est bien définie.

Exemple de stratification sur le sexe grade et la spécialité des enseignants d'éducation physique:

Effectifs de l'univers
Original

effectifs de l'échantillon
stratifié

Première variable : le grade


- 600 maîtres qui enseignent dans les écoles primaires (MP).....	60
- 400 professeurs qui enseignent dans les écoles secondaires (PS)	40
Total ou T = 1000.....	T = 100

Deuxième variable : le sexe

- 800 enseignants hommes (H).....	80
- 200 enseignantes femmes (F)	20
Total ou T = 1000	T' = 100

Troisième variable : la spécialité

- 500 spécialistes de sports individuels (SI) ..	.50
- 500 spécialistes de sports collectifs (SC)50
Total ou T = 1000	T' =100



Dans le **tableau 1**, le nombre 200 représente l'effectif de la **première strate**, c.-à-d. le nombre d'enseignants de sexe masculin ou hommes (H), qui sont des maitres (M), qui exercent dans les écoles primaires (P) et qui sont des spécialistes de sports individuels (SI). On voit que cette strate est le lieu de croisement des trois variables indépendantes :


Le **grade** (ce sont des maitres qui enseignent dans le primaire et non pas des professeurs qui enseignent dans le secondaire).

Le **sexe** (ce sont des hommes et non pas des femmes).

La **spécialité** (ce sont des spécialistes de sports individuels et non pas des spécialistes de sports collectifs).

Tab. 1: Description de l'univers en termes de strates

Sexe*Spécialité*Grade	M P	P S	Totaux
H * S I	200	170	370
H * S C	250	180	430
Sous-totaux	450	350	800
F*SI	100	30	130
F*SC	50	20	70
Sous-totaux	150	50	200
Totaux	600	400	1000



L'effectif de la seconde strate s'élève à 170: c'est le nombre des professeurs hommes (P^*H) qui enseignent dans le secondaire (S) et qui sont également spécialisés en sports individuels (SI).

Comme nous avons trois variables (le grade, le sexe et la spécialité) et comme chaque variable est dichotomique (maitre ou professeur pour ce qui est du grade, homme ou femme pour ce qui est du sexe et sports individuels ou collectifs pour ce qui est de la spécialité), nous devons obtenir $2 \times 2 \times 2 = 8$ strates (probabilité conditionnée).



Les six strates qui restent sont les suivantes :

- La troisième strate contient 250 maîtres spécialisés en sports collectifs.
- La quatrième contient 180 professeurs hommes spécialisés en sports collectifs.
- La cinquième est constituée de 100 institutrices spécialistes de sports individuels.


La sixième est constituée de 30 professeurs femmes spécialistes de sports individuels.

- La septième est composée de 50 institutrices spécialistes de sports collectifs.

Enfin, la huitième se compose de 20 professeurs femmes spécialisées en sports collectifs.

Tab. 2: Description de l'échantillon en termes de strates


Sexe*Spécialité*Grade	M P	P S	Totaux
H*SI	20	17	37
H*SC	25	18	43
Sous-totaux	45	35	80
F*SI	10	3	13
F*SC	5	2	7
Sous-totaux	15	5	20
Totaux	60	40	100



■ le tableau 2 qui représente l'échantillon est une reproduction fidèle (à une échelle réduite de 1/10 ou 10%) du tableau 1 dont il respecte proportionnellement la constitution, ce qui signifie que **l'échantillon stratifié est une image réduite de l'univers duquel il est extrait**. Le nombre des éléments est réduit, mais les rapports sont constants.

La méthode des quotas ou des contingents

Commencer tout d'abord, par stratifier l'univers, mais au lieu de tirer au hasard et à l'intérieur des strates les éléments qui vont constituer l'échantillon, laisser les enquêteurs libres d'interroger tout sujet qu'ils peuvent rencontrer pourvu que la condition **x pour cent** de chaque strate soit remplie, x étant la taille de l'échantillon exprimée en pourcentage par rapport à celle de l'univers.



■ Dans la méthode des quotas, chaque enquêteur se contente d'interroger les premiers venus qu'il rencontre et continue de la sorte tant que le quota n'est pas atteint, sans chercher à se déplacer très loin pour diversifier ses rencontres et ce en se conformant à la loi du "moindre effort." De cette manière, les enquêteurs risquent de ne solliciter que des personnes de même voisinage.

L'échantillonnage aréolaire

- Le sondage aréolaire s'impose lorsque **le chercheur ne dispose d'aucun renseignement à propos de l'univers**. Il ne dispose même pas d'un répertoire nominatif avec les adresses des éléments, répertoire qui constitue le minimum autorisant et justifiant le recours à la méthode au hasard simple. Dans cette situation, le chercheur ne connaît, en fait, que le **territoire géographique** que la population occupe.
- Dans une situation pareille, la seule solution qui lui reste consiste à appliquer **le principe du tirage au hasard en se basant, cette fois-ci, non pas sur les individus qui l'intéressent, mais sur le territoire géographique qu'ils occupent**, unique source de sa connaissance de la réalité.

- **Par conséquent, au lieu de tirer au sort les sujets dont l'identité lui échappe, il va tirer des zones du territoire qu'ils occupent.**
- **La méthode consiste à découper la surface géographique contenant la population en plusieurs aires ou zones de dimensions égales. On obtient alors une carte géographique quadrillée.**
- **On donne ensuite et comme d'habitude à chaque aire un matricule et on tire au sort, soit par hasard simple, soit par intervalles réguliers, un certain nombre de zones pour obtenir un échantillon de zones représentatif. Le chercheur interrogera enfin tous les occupants de chaque aire apparue au tirage.**