



Université Abderrahmane Mira de Bejaia
Faculté des sciences humaines et sociales
Département des sciences sociales

Intitulé du module
Statistique inferentielle

Destiné aux étudiants de première année

(Groupe:1-2-3-4-5)

Cours préparés par

Dr AMOUR-Mustapha

Année Universitaire : 2020-2021

Cours n° 1 : Echantillon et échantillonnage

1- Qu'est-ce qu'un échantillon, pourquoi un échantillon?

Un échantillon est constitué dès que l'on sélectionne un nombre restreint d'unités à partir d'une population d'unités. Cette population doit être définie de telle manière que l'on peut toujours savoir si une unité fait partie de la population.

Pourquoi échantillonner

- ✓ réduire les coûts.
- ✓ assurer un meilleur contrôle des opérations et une surveillance plus rigoureuse.

2- Définitions

- **La population:** est donc constituée de l'ensemble des unités auxquelles les résultats de l'enquête s'appliqueront.
- **Echantillon:** sous ensemble d'une population ou une partie de l'ensemble.
- **Echantillonnage:** opération consistant à choisir un sous ensemble de la population qui formera l'échantillon.
- **La base d'échantillonnage ou base de sondage:** est constituée par la liste des unités d'échantillonnage (liste matérielle ou conceptuelle), c'est-à-dire liste des unités à partir de laquelle se fera la sélection. Cette liste doit constituer la meilleure approximation possible de la population : *Chaque membre de la population doit y apparaître une fois et une seule fois.*

Exemple: Si je prends les diplômés de bac et doctorat des cinq dernières années, certaines personnes peuvent apparaître deux ou même trois fois dans la liste. En tant que diplômés de 1^{er}, 2^{ème} et 3^{ème} cycle; elles auront donc plus d'une chance d'être choisies; je devrais donc épurer la liste de toutes les doubles ou les triples entrées avant de tirer l'échantillon.

- **L'unité d'échantillonnage:** est constituée de chaque "membre" de la base d'échantillonnage. Elle renvoie au niveau auquel se fait l'échantillonnage; l'unité d'échantillonnage peut être un ménage (un appartement ou un numéro de téléphone), une école, un îlot, un étudiant.

A différencier de :

- **L'unité d'analyse :** le niveau sur lequel porte l'analyse; on sélectionne des ménages mais l'analyse porte sur des individus.
- **L'unité déclarante :** celle relative à l'informateur qui donne des informations sur d'autres personnes; par exemple, une mère interrogée sur ses enfants.
- **L'unité de référence:** celle sur laquelle on demande des informations; par exemple, les enfants pour lesquels la mère a donné des informations.

3- Les divers types d'échantillons:

Après avoir constitué la base de sondage, la deuxième étape consiste à déterminer la meilleure façon d'échantillonner. Plusieurs stratégies s'offrent à nous. Il faut déterminer laquelle sera la plus efficace tout en permettant de généraliser à la population.

On distingue deux catégories de méthodes d'échantillonnages

Méthodes aléatoires: (probabiliste)

Chaque unité a une chance que l'on peut quantifier d'être.

Méthodes non aléatoires: (non probabiliste)

A- L'échantillon probabiliste:

L'échantillonnage probabiliste entraîne la sélection d'un échantillon à partir d'une population, sélection qui repose sur le principe de la randomisation (la sélection au hasard ou aléatoire) ou la chance. Il est plus complexe, prend plus de temps et est habituellement plus coûteux que l'échantillonnage non probabiliste. Toutefois, comme les unités de la population sont sélectionnées au hasard et qu'il est possible de calculer la probabilité d'inclusion de chaque unité dans l'échantillon, on peut, grâce à l'échantillonnage probabiliste, produire des estimations fiables, de même que des estimations de l'erreur d'échantillonnage et faire des inférences au sujet de la population.

Il existe plusieurs méthodes différentes permettant de sélectionner un échantillon probabiliste. La méthode qu'on choisira dépendra d'un certain nombre de facteurs, comme la base de sondage dont on disposera, la façon dont la population sera distribuée, ce que sonder les membres de la population coûtera et la façon également dont les utilisateurs analyseront les données. Lorsque vous choisirez un plan d'échantillonnage probabiliste, votre but devrait consister à réduire le plus possible l'erreur d'échantillonnage des estimations pour les variables d'enquête les plus importantes, tout en réduisant le plus possible également le délai et le coût de réalisation de l'enquête.

Définition: technique d'échantillonnage, selon laquelle la probabilité qu'un individu quelconque fasse partie de l'échantillon est mesurable.

- Chaque unité a une chance que l'on peut quantifier d'être.

- Le tirage repose sur le hasard, la théorie des probabilités traite les phénomènes.

- Le tirage s'effectue de manière uniforme dans toute la population.

- Chaque unité doit avoir une probabilité connue d'être choisie. Cette probabilité ne peut pas être nulle mais elle n'est pas nécessairement égale pour toutes les unités.

Pourquoi utiliser un échantillon de type probabiliste?

- ✓ permet l'utilisation de méthodes d'estimation et de méthodes d'inférence et d'analyse statistique .
- ✓ toutes sont basées sur la théorie des probabilités.

Les méthodes aléatoires:

- L'échantillon aléatoire simple
- L'échantillonnage systématique
- L'échantillonnage stratifié
- L'échantillonnage en grappes

Cours n° 2

1- L'échantillonnage aléatoire simple:

Définition: technique d'échantillonnage, selon laquelle tous les individus ont la même probabilité de faire partie de l'échantillon.

- Dans un *échantillonnage aléatoire simple*, chaque membre d'une population a une chance égale d'être inclus à l'intérieur de l'échantillon. Chaque combinaison de membres de la population a aussi une chance égale de composer l'échantillon. Ces deux propriétés sont ce qui définit un échantillonnage aléatoire simple. Vous devez dresser une liste de toutes les unités incluses dans la population observée pour sélectionner un échantillon aléatoire simple.
- consiste à choisir des individus de telle sorte que chaque membre de la population a une chance égale de figurer dans l'échantillon.

Ce choix peut se faire avec remise ou sans remise :

Avec remise: un individu peut être choisi plusieurs fois.

Sans remise: un individu déjà choisi ne peut l'être de nouveau. (C'est le cas habituel).

- ✓ **Avantage de cette méthode :** On peut espérer un échantillon «représentatif » puisque la méthode donne à chaque individu de la population une chance égale.
- ✓ **Difficultés :** la méthode n'est applicable que lorsqu'il existe une liste exhaustive de toute la population.

Comment procéder ?

- On numérote tous les individus de la liste correspondant aux individus de la population avec des nombres comportant un même nombre de chiffres.
- En utilisant une table de nombres aléatoires, une calculatrice ou un programme informatique (BASIC, PASCAL, SPSS, etc.), on obtient des nombres aléatoires comportant le nombre de chiffres désiré.
- On sélectionne les nombres qui coïncident avec la liste. On rejette les nombres qui ne coïncident pas avec la liste ou qui se répètent, on s'arrête après avoir sélectionné n individus. (n représentant le nombre d'individus souhaités dans l'échantillon).
- **Exemple n° 1 :**

Il faudrait numéroté dans un ordre séquentiel chaque entrée ou inscription pour prélever un échantillon aléatoire simple d'un annuaire téléphonique. S'il y avait 10 000 entrées dans l'annuaire téléphonique et si la taille de l'échantillon était 2 000 numéros, un ordinateur devrait alors générer au hasard 2 000 numéros entre 1 et 10 000. Chaque numéro aurait la même chance qu'un autre d'être généré par l'ordinateur (ce qui respecterait l'exigence de l'échantillonnage aléatoire simple : une chance égale pour chaque unité). Les 2 000 entrées dans l'annuaire téléphonique correspondant aux 2 000 numéros aléatoires générés par l'ordinateur composeraient l'échantillon.

Un tirage à la loterie est un bon exemple d'échantillonnage aléatoire simple. Par exemple, lorsqu'un échantillon de six numéros est généré au hasard à partir d'une

population de 49 numéros, chacun de ces derniers à une chance égale d'être sélectionné et chaque combinaison de six numéros a la même chance d'être la combinaison gagnante. Même si les gens tendent à éviter une combinaison comme 1-2-3-4-5-6, cette combinaison a la même chance d'être la série gagnante de numéros que la combinaison 8-15-21-28-32-40.

➤ **Exemple n° 2 :**

Supposez que votre école compte 500 élèves et que vous devez mener une courte enquête sur la qualité des aliments servis à sa cafétéria. Vous déterminez qu'un échantillon de 10 élèves devrait suffire à vos fins. Pour obtenir votre échantillon, vous attribuez à chaque élève de votre école un numéro compris entre 1 et 500. Pour sélectionner cet échantillon, vous utilisez une table de numéros générés au hasard. Tout ce que vous avez à faire consiste à prendre un point de départ à l'intérieur de la table (un numéro de rangée et un numéro de colonne) et à examiner les numéros aléatoires qui y figurent. Dans ce cas, puisque les données ne dépassent pas trois chiffres, les numéros aléatoires devraient renfermer trois chiffres également. Ne tenez pas compte des numéros aléatoires supérieurs à 500, parce qu'ils ne correspondent à aucun des élèves de votre école. Rappelez-vous que votre échantillon est un échantillon sans remplacement et que, si un numéro se répète, vous devez le sauter et utiliser le numéro aléatoire suivant. Les 10 premiers numéros différents entre 001 et 500 composent votre échantillon.

2- L'échantillonnage systématique:

- **Définition:** technique d'échantillonnage consistant à choisir les individus d'un échantillon au moyen d'un poids de sondage $K = N/n$.
- Parfois appelé *échantillonnage par intervalles*, l'*échantillonnage systématique* signifie qu'il existe un écart, ou un intervalle, entre chaque unité sélectionnée qui est incluse dans l'échantillon. Vous devez suivre les étapes énumérées ci-dessous pour sélectionner un échantillon systématique :
 - 1 Numéroté de 1 à **N** les unités incluses dans votre base de sondage (où **N** est la taille de la population totale).
 - ✓ Déterminer l'intervalle d'échantillonnage (**K**) en divisant le nombre d'unités incluses dans la population par la taille de l'échantillon que vous désirez obtenir. Par exemple, pour sélectionner un échantillon de 100 unités à partir d'une population de 400, vous auriez besoin d'un intervalle d'échantillonnage de $400 \div 100 = 4$. $K = 4$, par conséquent. Vous devrez sélectionner une unité sur 4 pour avoir finalement au total 100 unités à l'intérieur de votre échantillon.
 - ✓ Sélectionner au hasard un nombre entre 1 et **K**. Ce nombre s'appelle l'origine choisie au hasard et serait le premier nombre inclus dans votre échantillon. À l'aide de l'échantillon fourni ci-dessus, vous sélectionneriez un chiffre entre 1 et 4 à partir d'une table de nombres aléatoires (pris au hasard). Si vous choisissiez 3, la troisième unité incluse dans votre base de sondage serait la première

unité comprise dans votre échantillon; si vous choisissiez 2, le début de votre échantillon serait la deuxième unité incluse dans votre base de sondage.

- ✓ Sélectionner chaque K^e (dans ce cas, chaque 4^e) unité après ce premier nombre. L'échantillon pourrait, par exemple, se composer des unités suivantes de façon à constituer un échantillon de 100 : 3 (l'origine choisie au hasard), 7, 11, 15, 19... 395, 399 (jusqu'à N , qui est 400 dans ce cas).

Vous pouvez constater, à l'aide de l'exemple fourni ci-dessus, que dans le cas d'un échantillonnage systématique, seuls quatre échantillons possibles, qui correspondent aux quatre origines choisies au hasard également possibles, peuvent être sélectionnés:

1, 5, 9, 13... 393, 397

2, 6, 10, 14... 394, 398

3, 7, 11, 15... 395, 399

4, 8, 12, 16... 396, 400

Chaque membre de la population ne fait partie que de l'un des quatre échantillons et chaque échantillon a une chance égale d'être sélectionné. Cela nous permet de constater que chaque unité a une chance sur quatre d'être sélectionnée à l'intérieur de l'échantillon. Sa probabilité d'être sélectionnée est la même que si l'on sélectionnait un échantillon aléatoire simple de 100 unités. La principale différence tient au fait que dans le cas d'un échantillonnage aléatoire simple, toute combinaison de 100 unités aurait une chance de constituer l'échantillon, tandis que dans celui d'un échantillonnage systématique, il n'y a que quatre échantillons possibles. Cela nous permet aussi de constater à quel point l'échantillonnage systématique est précis comparativement à l'échantillonnage aléatoire simple. L'ordre de la population incluse dans la base de sondage déterminera les échantillons possibles pour l'échantillonnage systématique. Si la population est distribuée au hasard dans la base de sondage, un échantillonnage systématique devrait alors produire des résultats similaires à ceux d'un échantillonnage aléatoire simple.

- **Avantages** : facile à sélectionner parce qu'un seul individu est choisi au hasard. On peut obtenir une bonne précision parce que la méthode permet de répartir l'échantillon dans l'ensemble de la liste.
- **Désavantages** : Les données peuvent être biaisées à cause de la périodicité.

▪ **Exemple n° 1:**

Imaginez que vous devez mener une enquête pour votre collège ou votre université sur les logements pour les étudiants. Dix mille (10 000) étudiants sont inscrits dans votre établissement d'enseignement et vous voulez en prélever un échantillon systématique de 500. Pour ce faire, vous devez premièrement déterminer ce que serait votre intervalle d'échantillonnage (K) :

- ✓ Population totale ÷ taille de l'échantillon = intervalle d'échantillonnage
$$N \div n = K$$
$$= 10\,000 \div 500$$
$$= 20$$
- ✓ Il faudrait attribuer un numéro séquentiel à chaque étudiant pour entreprendre cet échantillonnage systématique. On choisirait le point de départ en sélectionnant un numéro au hasard entre 1 et 20. Si ce numéro était 9, on sélectionnerait alors le 9e étudiant inscrit sur la liste et chaque 20e étudiant par la suite. L'échantillon d'étudiants serait constitué de ceux qui correspondraient aux numéros d'étudiant 9, 29, 49, 69... 9 929, 9 949, 9 969 et 9 989.
- **Exemple n° 2 :**
 - ✓ Supposez que vous dirigez une épicerie de grande surface et que vous possédez une liste des employés de chacune de ses sections. L'épicerie est divisée entre les 10 sections suivantes : le comptoir de charcuterie, la boulangerie, les caisses, les stocks, le comptoir des viandes, les fruits et légumes, la pharmacie, le magasin de photographie, le magasin de fleurs et le nettoyage à sec. Chaque section compte 10 employés, y compris un gérant (ce qui fait 100 employés au total). Votre liste est ordonnée par section, le gérant y étant énuméré le premier et les autres employés y étant ensuite inscrits dans l'ordre décroissant d'ancienneté.
 - ✓ Si vous voulez sonder vos employés au sujet de leurs réflexions sur leur milieu de travail, vous pourriez choisir un petit échantillon pour répondre à vos questions. Si vous utilisiez un échantillonnage systématique et si votre intervalle d'échantillonnage était 10, vous pourriez alors ne sélectionner finalement que les gérants ou que les employés de chaque section ayant le moins d'ancienneté. Ce type d'échantillon ne vous donnerait pas un portrait complet ni approprié des réflexions de vos employés.
- Si la population est distribuée au hasard dans la base de sondage, un échantillonnage systématique donnera des résultats similaires à ceux d'un échantillonnage aléatoire simple.

Cours n° 3

3-L'échantillonnage stratifié:

- **Définition:** technique d'échantillonnage consistant à subdiviser la population en sous groupes ou strates relativement homogènes, puis à prélever dans chaque strate un échantillon aléatoire simple.
 - ✓ Lorsqu'on utilise l'*échantillonnage stratifié*, on divise la population en groupes homogènes (appelés strates), qui sont mutuellement exclusifs, puis on sélectionne à partir de chaque strate des échantillons indépendants. On peut utiliser n'importe quelle des méthodes d'échantillonnage mentionnées dans la présente section (et il en existe d'autres) pour sélectionner l'échantillon à l'intérieur de chaque strate. La méthode d'échantillonnage peut varier d'une strate à une autre. Lorsqu'on utilise l'échantillonnage aléatoire simple pour sélectionner l'échantillon à l'intérieur de chaque strate, on appelle le plan d'échantillonnage un plan d'échantillonnage aléatoire simple stratifié. On peut stratifier avant l'échantillonnage une population au moyen de toute variable dont on dispose pour la totalité des unités incluses dans la base de sondage (comme l'âge, le sexe, la province de résidence, le revenu, etc.)
 - ✓ L'échantillonnage stratifié nous assure d'obtenir une taille d'échantillon suffisante pour des sous-groupes de la population à laquelle nous nous intéressons. Étant donné que chaque strate devient une population indépendante lorsque vous stratifiez une population, vous devrez déterminer pour chaque strate la taille de l'échantillon.
- Démarche de sélection :
 - 1) On subdivise la population en strates (groupes relativement homogènes) qui sont mutuellement exclusives.
 - 2) Proportionnellement à son importance dans la population, on calcule combien il faut d'individus au sein de l'échantillon pour représenter chaque strate.
 - 3) Dans chacune des strates, on choisit au hasard le nombre nécessaire d'individus.Les variables de stratification doivent être :
 - ✓ Simple à utiliser.
 - ✓ Facile à observer.
 - ✓ Étroitement reliées au thème de l'enquête.
- **Avantages :** Il est peu probable de choisir un échantillon absurde puisqu'on s'assure de la présence proportionnelle de tous les divers sous-groupes composant la population.
- **Désavantages :** La méthode suppose l'existence d'une liste de la population. Il faut aussi connaître comment cette population se répartit selon certaines strates.

▪ **Exemple n° 1 :**

Un conseil scolaire voulait évaluer l'opinion des élèves sur les programmes de la 3^e année d'études secondaires. Il a décidé de sonder les élèves de l'école secondaire de ville et a utilisé une technique d'échantillonnage stratifié pour s'assurer de sélectionner un échantillon représentatif d'élèves de toutes les années d'études offertes dans cet établissement d'enseignement.

Dans ce cas, les strates étaient les cinq années d'études (les 1, 2 et 3^e années). Le conseil scolaire a ensuite sélectionné un échantillon à l'intérieur de chaque strate. Les noms des élèves sélectionnés à l'intérieur de cet échantillon en ont été extraits à l'aide de la méthode d'échantillonnage aléatoire simple ou d'échantillonnage systématique, ce qui a donné un échantillon total de 100 élèves.

▪ **Exemple n° 2 :**

Le tableau suivant nous donne la répartition des médecins d'une clinique selon leurs spécialités.

Spécialités	Généralistes	cardiologues	Pédiatres	Ensembles
Effectifs	51	35	22	108

On souhaite prélever un échantillon de 42 médecins pour les interroger sur la communication avec les patients.

Questions: Calculer les effectifs de cet échantillon en utilisant la méthode d'échantillonnage stratifié.

Spécialités	Généralistes	cardiologues	Pédiatres	Ensembles
Effectifs	51	35	22	108
%	47,22	32,40	20,37	100
Echantillon	20	14	8	42

La stratification est des plus utiles lorsque les variables de stratification sont :

- simples à utiliser;
- faciles à observer;
- étroitement reliées au thème de l'enquête.

4-L'échantillonnage par grappes:

- **Définition:** technique d'échantillonnage consistant à définir des sous-groupes naturels de la population, appelé grappes, à choisir des grappes au hasard et à effectuer des mesures pour tous les individus de chaque grappe sélectionnées.
 - ✓ Il est parfois trop dispendieux de disséminer un échantillon dans l'ensemble de la population. Les coûts de déplacement risquent de devenir élevés lorsque les intervieweurs doivent sonder des gens d'un bout à l'autre du pays. Les statisticiens peuvent choisir la technique de l'échantillonnage en grappes pour réduire les coûts.
 - ✓ La technique de l'échantillonnage en grappes entraîne la division de la population en groupes ou en grappes comme son nom l'indique. Suivant cette technique, on sélectionne au hasard un certain nombre de grappes pour représenter la population totale, puis on englobe dans l'échantillon toutes les

unités incluses à l'intérieur des grappes sélectionnées. On n'inclut dans l'échantillon aucune unité de grappes non sélectionnées; ces unités sont représentées par celles tirées de grappes sélectionnées. La technique en question diffère de la technique d'échantillonnage stratifié, qui entraîne la sélection d'unités de chaque groupe.

- ✓ Mentionnons, entre autres exemples de grappes, les usines, les établissements d'enseignement et les régions géographiques telles que les subdivisions électorales. On utilise les grappes sélectionnées pour représenter la population.
- ✓ Dans les méthodes précédentes, l'unité statistique était choisie individuellement.
- ✓ La technique de l'échantillonnage en grappes entraîne la division de la population en groupes ou grappes.
- ✓ On sélectionne au hasard un certain nombre de grappes (unités primaires) pour représenter la population.
- ✓ On sélectionne tous les individus des grappes choisies.
- **Avantages** : la méthode ne nécessite pas une liste globale de la population puisque seules les individus inclus dans les grappes comptent. Elle permet de limiter l'échantillon à des groupes compacts ce qui permet de réduire les coûts de déplacement, de suivi et de supervision.
- **Désavantage** : la méthode peut entraîner des résultats imprécis (moins précis que les méthodes précédentes) puisque les unités voisines ont tendance à se rassembler. Elle ne permet pas de contrôler la taille finale de l'échantillon.
- **Exemple n° 1:**

afin de réaliser une étude de marché sur un nouveau produit auprès des consommateurs, on subdivise une ville en quartier et chaque quartiers en rues, puis on choisi de manière aléatoire quartes(04) rues dans différents quartiers de la ville . Chacune des rues choisie (retenus) constitue une grappes, On interroge ensuite toute les personnes résidentes dans l'une des rues choisie.
- **Exemple n° 2 :**

Supposez que vous représentez une organisation d'athlétisme désirant déterminer quels sports pratiquent les élèves de 1^{er} année en Algérie. Il serait trop coûteux et trop long d'interroger chaque élève Algérien de 1^{er} année ou même deux ou trois élèves de chaque classe de 1^{er} année en Algérie. On sélectionne plutôt au hasard 100 écoles de tout le pays.

Ces écoles fournissent des grappes d'échantillons. On sonde ensuite chaque élève de 1^eannée de chacune des 100 grappes. Les élèves inclus dans ces grappes représentent, en effet, tous les élèves de 1^{er} année en Algérie.

Exercice: sur la méthode d'échantillonnages probabiliste

On a une liste d'élèves comprenant $K = 40$ élèves et on désire en choisir $k = 10$ élèves de façon aléatoire simple.

1. On numérote de 1 à K , c'est-à-dire de 1 à 40 les prénoms des élèves, ce qui donne :

1. Lyès	9. Sannaa	17. Shahinda	25. Wassim	33. Mokhtar
2. Hussam	10. Walid	18. Maya	26. Younes	34. Mustapha
3. Chouaib	11. Syrine	19. Katia	27. Racha	35. Ryma
4. Kawther	12. Hichem	20. Mourad	28. Kamel	36. Amira
5. Lyna	13. Ahmed	21. Kahina	29. Lakhder	37. Nadine
6. Farouk	14. Fatma	22. Aymen	30. Fouad	38. Djawad
7. Meriem	15. Amel	23. Rachid	31. Khaled	39. Amira
8. Farah	16. Hanane	24. Nasser	32. Saida	40. Farid

2. On sélectionne de façon aléatoire $k = 10$ numéros entre 1 et 40 inclusivement, c'est-à-dire qu'on sélectionne :

- un premier numéro entre 1 et 40, ce qui donne par exemple 16 ;
- un deuxième numéro entre 1 et 40, différent du précédent, et ainsi de suite jusqu'à l'obtention de 10 numéros différents, ce qui donne par exemple la série : 16, 22, 30, 12, 29, 20, 13, 18, 26 et 8.

3. On constitue l'échantillon en se retenant l'élève correspondant à chaque numéro, en se référant à ces 10 numéros cités au dessus, quels sont les prénoms des élèves sélectionnés ?

Solution : Les prénoms sélectionnés : 16 Hanane, 22 Aymen, 30 Fouad, 12 Hichem, 29 Lakhder, 20 Mourad, 13 Ahmed, 18 Maya, 26 Younes, 8 Farah.

4. On désire en choisir $n = 10$ élèves de façon aléatoire systématique, on détermine le pas de sondage, désigné par k , donné par : $k = N/n$, on prend k comme point de départ, quels sont les 10 élèves sélectionnés pour l'échantillon ?

Solution : $k = 40/10 = 4$ donc le premier sélectionné c'est le prénom qui porte le numéro 4 : kawther, ensuite qui porte le numéro 8 c'est : Farah, le numéro 12 Hichem, le numéro 16 Hanane, le numéro 20 Mourad, le numéro 24 Nasser, le numéro 28 Kamel, le numéro 32 Saida, le numéro 36 Amira, le numéro 40 Farid.

5. Le pas de sondage étant égal à 4, on détermine le point de départ, désigné par d , en choisissant au hasard un nombre entre 1 et 4, ce qui donnerait par exemple 2. Constituez l'échantillon en retenant l'élève correspondant à chaque numéro ?

Solution : c'est les élèves qui porte le numéro 2 6 10 14 18 22 26 30 34 38 .

6. On désire établir un échantillon par grappe constitué de 08 élèves, combien de grappes peut-on construire ?

Solution : en peut construire 05 grappes : $40G = 40/8 = 5$

7. Etablissez un échantillon stratifié par sexe de 10 élèves sachant qu'on a 60 filles et 40 garçons ?

Solution : par strate (sexe)

Nombre de garçons = $40 * 10 / 100 = 4$ garçons

Nombre de filles = $60 * 10 / 100 = 6$ filles

Cours n° 4**B-L'échantillon non probabiliste:****Méthodes empirique ou non aléatoires**

- **Définition:** technique d'échantillonnage dans laquelle le choix des éléments est arbitraire et la probabilité qu'un élément fasse partie de l'échantillon est inconnu.
- ✓ Un échantillon non probabiliste est un échantillon qui n'offre pas à tous les membres de la population une chance égale, d'être sélectionnés la probabilité de sélection d'un membre de la population est donc inconnue, il devient alors impossible de calculer la précision des résultats ainsi obtenu et d'utiliser les résultats pour extrapoler sur l'ensemble de la population. Cette impossibilité réside essentiellement dans le fait qu'il est impossible que les répondants peuvent ne pas être représentatifs de la population.
- ✓ On oppose aux méthodes aléatoires les méthodes non aléatoires.
- ✓ Les méthodes non aléatoires sont des méthodes où le concept de «chance égale» est absent. Ce sont des méthodes généralement peu fiables.
- ✓ Elles ne nécessitent pas de base de sondage
- ✓ Elles sont souvent utilisées
 - pour des études exploratoires;
 - pour réduire les coûts;
 - quand il est impossible ou non envisageable d'utiliser la méthode aléatoire.

On distingue :

- **L'échantillonnage à l'aveuglette ou de commodité** : Ex: Déguster un échantillon de l'huile d'olive.....
- **L'échantillonnage de volontaires** : Ex : Expériences médicales ou psychologiques.
- **L'échantillonnage au jugé** : cette méthode implique la sélection d'individus en fonction de l'idée qu'on se fait de la composition de la population. On le fait pour des essais auprès des groupes cibles.
- **L'échantillonnage par quotas** : il est largement utilisé dans les enquêtes d'opinion et les études de marché notamment parce qu'il ne suppose pas de liste des individus de la population. On parle aussi **d'échantillonnage dirigé ou par choix raisonné**. On demande aux enquêteurs de faire un nombre d'entrevues dans divers groupes établis en fonction du secteur géographique, de l'âge, du sexe ou d'autres caractéristiques... L'enquêteur doit respecter son quota.
- **Avantages** : Moins coûteuse et plus facile à réaliser.
- **Désavantages**: Beaucoup de non-réponses; difficulté de trancher lorsqu'il s'agit de sélectionner des individus d'un groupe d'âge ouvert (Ex : 65 ans et plus : faut-il prendre 66 ans, 70 ans ...)

1- Echantillonnage accidentel

- **Définition:** technique d'échantillonnage non aléatoire consistant à prélever un échantillon de façon totalement arbitraire.
Il s'agit d'un échantillon constitué d'individus qui se trouvaient accidentellement à l'endroit et au moment où l'information a été collectée.
Les échantillons accidentels ne peuvent être considérés représentatifs d'aucune population. Il est risqué de généraliser à une population donnée des résultats obtenus par un échantillon accidentel.

Exemple :

- -Enquêtes réalisées dans la rue, les lieux publics, en sortie de super marché ...
- Questionnaires figurant dans les magazines et renvoyés spontanément.

2- Echantillonnage à priori:

- ✓ C'est un échantillonnage par jugement à priori. Il consiste à sélectionner des individus dont on pense, avant de les interroger, qu'ils peuvent détenir l'information.
- ✓ Le risque de ce type d'échantillonnage est de considérer des individus, apparemment représentatifs de la population étudiée.

3- Echantillonnage Boule de neige :

Définition: il suffit de connaître quelques éléments de la population ciblée, grâce auquel on joint les autres individus.

- ✓ Cette méthode est réservée aux populations composées d'individus dont l'identification est difficile ou qui possèdent des caractéristiques rares.
- ✓ La méthode consiste à faire construire l'échantillon par les individus eux-mêmes. Il suffit d'en identifier un petit nombre initial et de leur demander de faire appel à d'autres individus possédant les mêmes caractéristiques.

4- Echantillonnage par Quotas:

Définition: technique d'échantillonnage non aléatoire consistant à décomposer la population en sous groupes et à laisser l'enquêteur choisir un nombre déterminé d'individu dans chaque sous-groupe.

- ✓ L'échantillonnage par quotas est l'échantillonnage non probabiliste le plus connu, et finalement le mieux accepté comme substitut aux méthodes probabilistes dans le cas où ces dernières rencontreraient des contraintes de base de sondage. Mais la représentativité de la population étudiée reste douteuse.
- ✓ L'échantillonnage par quotas consiste à étudier la structure de la population selon des critères choisis (quotas) empiriquement. L'échantillon est ensuite construit de manière à constituer une reproduction en miniature de la population sur ces critères.
- ✓ L'échantillonnage par quotas est une forme simplifiée de l'échantillonnage stratifié à fraction de sondage constante. Les quotas représentent les variables de stratification.
- ✓ Une fois les quotas sont fixés, les individus sont sélectionnés à la convenance de l'enquêteur.
- ✓ Les critères servant de base à la définition des quotas ne doivent pas être nombreux. Au-delà de 3 critères, la démarche devient complexe. Les quotas doivent être

construits sur une base de données fiables (statistiques disponibles) indiquant la répartition de la population sur les critères choisis. Les critères les plus utilisés dans les études de marché sont économiques et socio- démographiques en particulier l'âge, le sexe, la catégorie socioprofessionnelle.

Exemple:

Si on part d'une population ,telle que décrite dans le tableau suivant:

Caractéristiques	Effectifs
<u>Selon le genre</u>	
Masculin	120
Féminin	80
<u>Selon l'âge</u>	
25	100
30	62
50	38
<u>Selon le milieu géographique</u>	
Urbain	135
Rural	65

Question: déterminer l'échantillon de 83 individus par la méthode des quotas?

Caractéristiques	Effectifs	%	Echantillon
<u>Selon le genre</u>			
Masculin	120	60	50
Féminin	80	40	33
<u>Selon l'âge</u>			
25	100	50	41
30	62	31	26
50	38	19	16
<u>Selon le milieu géographique</u>			
Urbain	135	67,5	56
Rural	65	32,5	27

⊙ Démonstration:

➤ On calcule d'abord les pourcentages de chaque variable.

$$\text{Selon le genre: } \frac{120 \times 100}{200} = 60\% \quad \frac{80 \times 100}{200} = 40\%$$

$$\text{Selon l'âge: } \frac{100 \times 100}{200} = 50\% \quad \frac{62 \times 100}{200} = 31\% \quad \frac{38 \times 100}{200} = 19\%$$

$$\text{Selon le milieu géographique: } \frac{135 \times 100}{200} = 67,5\% \quad \frac{65 \times 100}{200} = 32,5\%$$

➤ On calcule l'échantillon de chaque variable.

$$\text{Selon le genre: } \frac{60 \times 83}{100} = 50 \quad \frac{40 \times 83}{100} = 33$$

$$\begin{array}{l} \text{Selon l'âge:} \quad \frac{50 \times 83}{100} = 41 \quad \frac{31 \times 83}{100} = 26 \quad \frac{19 \times 83}{100} = 16 \\ \text{Selon le milieu géographique:} \quad \frac{67,5 \times 83}{100} = 56 \quad \frac{32,5 \times 83}{100} = 27 \end{array}$$

Série n° 1 : Méthodes d'échantillonnage**Exercice 1:**

On a une liste comprenant **36** élèves et on désire choisir **10** élèves. Constituer votre échantillon en utilisant une méthode

- 1- Aléatoire simple.
- 2- Systématique.
- 3- Stratifiée selon le sexe.
- 4- Par grappes, en constituant 9 grappes de 4 élèves et prélever un échantillon de $n = 12$ élèves.
- 5- Par quotas selon l'âge, le sexe, et la classe.

Tableau 1 : Liste nominative des élèves selon l'âge et la classe.

Prénom	Age	Classe	Prénom	Age	Classe
Katia	16	1	Sihem	16	2
Fairouz	17	1	Raouf	16	1
Salim	17	2	Fahima	19	2
Ryma	18	2	Mourad	17	1
Kahina	18	2	Houda	19	2
Nacer	17	1	Bilal	17	2
Zakia	19	2	Hana	17	2
Yassine	17	2	Nassim	17	1
Rafik	17	2	Amine	18	2
Nadjet	16	1	Soria	17	1
Lamia	18	1	Zoubir	18	2
Lamine	19	2	Thanina	17	2
Nourhane	18	2	Kamelia	18	1
Reda	18	2	Djamel	17	2
Baya	19	1	Nesrine	17	1
Said	17	1	Ahmed	17	2
Loubna	18	1	Aymen	17	1
Katia	19	1	Sabrina	18	1

Exercice 2:

Un hôpital à **1820** personnel médical. Le directeur décide de demander aux infirmiers des suggestions afin d'améliorer leur condition de travail au sein de leur service. Le responsable choisit d'échantillonner systématiquement **140** infirmiers.

- 1- Quelle serait l'intervalle d'échantillonnage?
- 2- Si le numéro **14** était votre premier numéro tiré au hasard, quels seraient les **09** premiers numéros de votre échantillon ?

Exercice 3 :

Le tableau suivant donne la répartition des médecins d'une clinique privée.

Spécialités	Généralités	Cardiologues	Pédiatres	Ensemble
Effectif	102	72	43	217

On souhaite prélever un échantillon de **65** médecins pour les interroger sur la communication avec les patients.

Question : Calculer les effectifs de cet échantillon en utilisant la méthode d'échantillonnage stratifié.

Exercice 4 :

On a distribué un questionnaire sur une population de **202** étudiants, soit **69** étudiants en **1^{er}** année, **35** de **2^{ème}**, **28** de **3^{ème}**, **24** de **4^{ème}**, et **46** de **5^{ème}**. **105** étaient des garçons et les **97** autres étaient des filles. Trente- cinq (**35**) étudiants ont rempli le questionnaire.

Questions :

a-Que représente ces **35** étudiants ?

b-Sachant que l'équipe de cette recherche a procédé à un échantillonnage pour généraliser ces résultats Déterminer la structure de l'échantillon sondé par la méthode stratifié ?

Exercice 5 :

Voici les caractéristiques de **172** patients de 02 hôpitaux dont on connaît le sexe, le niveau et types de maladie. Il s'agit de déterminer pour cette population parenté de **65** patients en vue d'une enquête sur leurs appréciations des rapports entre médecins et patients.

Caractéristiques	Effectifs de la population
<u>Selon le sexe</u>	
Masculin	96
Féminin	76
<u>Selon le niveau</u>	
Bon	41
Moyen	85
Faible	46
<u>Type de maladie</u>	
Diabète	48
Hypertension	55
Gastrique	69

Question: Déterminer l'échantillon de patients par la méthode des quotas?