

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



Université Ammar Telidji Laghouat

Faculté des Sciences

Département de sciences agronomiques

Filière: Sciences alimentaires

Spécialité: Technologies agroalimentaires et contrôle de qualité

Module de :

MÉTHODES D'ÉCHANTILLONNAGE

*Cours destinés aux étudiants 3^{ème} année licence
Technologies agroalimentaires et contrôle de qualité*

Présenté par

Dr. ALLALI Khadidja

PLAN DE COURS:

Méthodes d'échantillonnage



1. Informations sur le cours

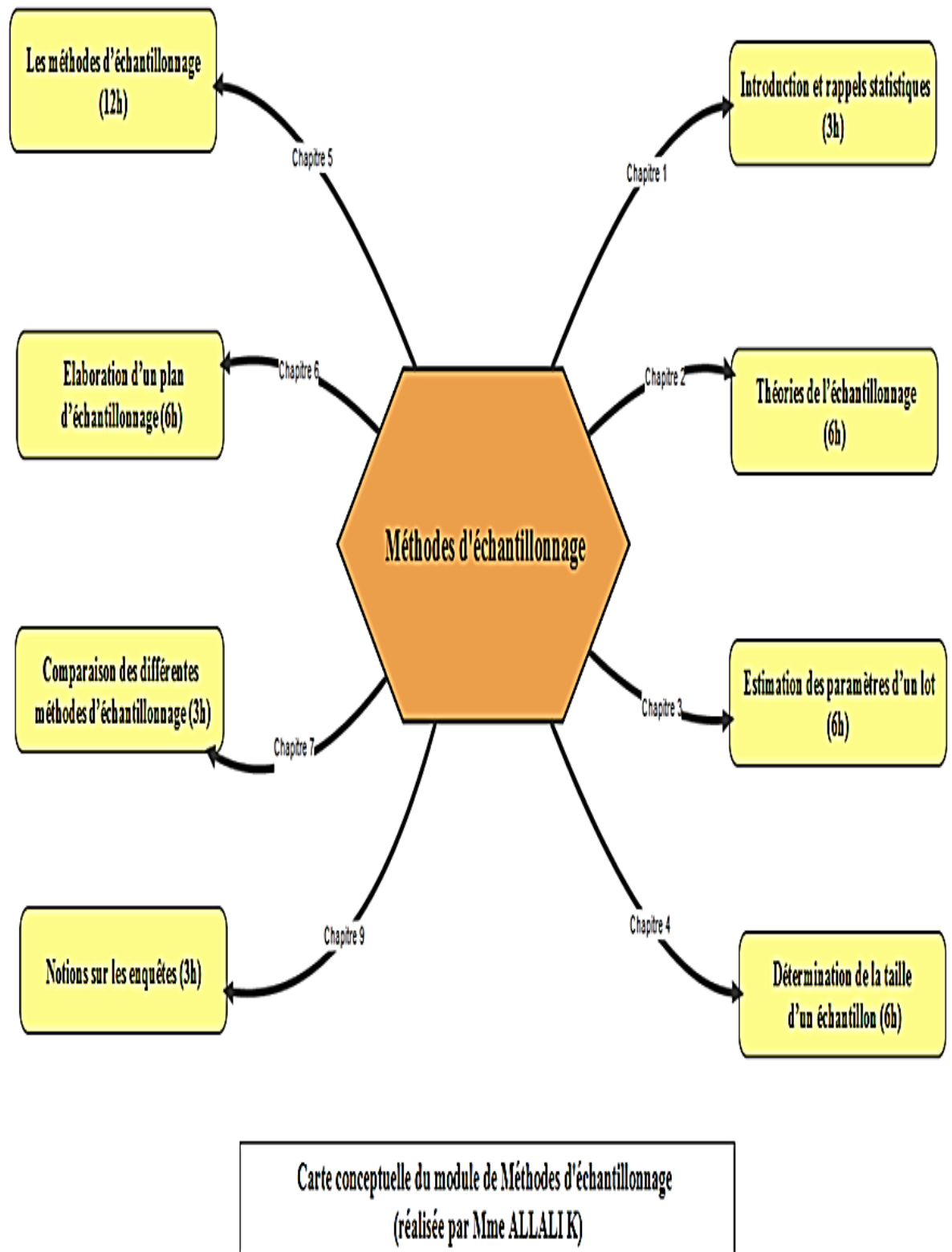
Faculté: Sciences
Département: Sciences agronomiques
Filière: Sciences alimentaires
Spécialité: Technologies agroalimentaires et contrôle de qualité
Intitulé du cours: Méthodes d'échantillonnage (U.E. Méthodologique (M5.1))
Crédit: 04
Coefficient: 02
Durée : 45 heures
Horaire et salle: Dimanche: 08h00-09h30. Amphi: A/Salle 1.7
Enseignant: Cours/TD/TP: Mme ALLALI Khadidja Contact: par mail à l'adresse allalikhadidjaagro@gmail.com
Disponibilité : Au bureau: Dimanche de 09h30 -13h00, Mardi de 9h00-14h00 Réponses sur le forum : toute question en relation avec le cours doit être postée sur le forum dédié (MOODLE) pour que vous puissiez, tous, tirer profit de ma réponse, je m'engage à répondre aux questions postées dans un délai de 72 heures. Par mail: réservé aux urgences, je m'engage à répondre immédiatement dans la mesure du possible.

2. Présentation du cours

Ce module traite les différentes méthodes utilisées pour faire l'échantillonnage des denrées alimentaires pour des études antérieures comme le contrôle microbiologique, dosage de sucres, etc.

Les cours de Méthodes d'échantillonnage sont destinés aux étudiants de 3^{ème} année licence. Dont l'objectif principal est de permettre aux étudiants de consolider leurs connaissances en biostatistiques et l'échantillonnage des denrées alimentaires afin de développer leur capacité à comprendre des concepts clés.

Ce cours de Méthodes d'échantillonnage présente en huit chapitres. La carte conceptuelle ci-dessous simplifiée du module vous donnera la manière donc nous allons procéder pour l'étude de ces fléaux.



3. Modalités de fonctionnement

Le cours est organisé en :

- Séances théoriques afin de vous transmettre l'ensemble des savoirs permettant de cerner rapidement les définitions de bases et les méthodes de travail ;
- Séances de travaux dirigées (TD), afin que vous puissiez mobiliser les savoirs dans la résolution des exercices ;
- Séances de travaux pratiques, afin de mobiliser les savoirs acquis dans la conception de la détermination de la taille de l'échantillon, le nombre des échantillons prélevés, la méthode d'échantillonnage, etc. ;
- Le déroulement du cours est assuré en **hybride**, en présentiel (**Amphi**) et à distance via la plateforme d'enseignement à distance de l'université de Laghouat (**MOODLE**). Cela permettra aux étudiants de revoir ou d'approfondir les concepts vus en présentiel. Les étudiants peuvent travailler collectivement dans la résolution des exercices de TD sous mon encadrement.

4. Pré-requis

Avant d'aborder ce module l'apprenant doit avoir des notions sur :

- Les connaissances de base en statistiques ;
- Les compétences sur l'utilisation d'outils informatique : maîtrise de l'Excel, Word, etc.

5. Visées d'apprentissage

L'objectif essentiel de l'enseignement de ce module est de comprendre l'importance de la représentativité de l'échantillon et assimilation des principales méthodes d'échantillonnage appliquées aux aliments.

Le cours de Méthodes d'échantillonnage vise à :

- 1)** En termes de connaissances, à apprendre les notions de base en statistiques, les théories d'échantillonnage, estimation des paramètres d'un lot et les différentes méthodes d'échantillonnage ;

- 2) En termes de savoir-faire, à entraîner les étudiants de savoir mettre en œuvre un échantillonnage représentatif ;
- 3) En termes de savoir-être, à orienter les étudiants vers l'éveil de la curiosité et vers l'application de cet enseignement dans leur domaine.

6. Modalités d'évaluation

Note finale = Un examen semestriel (60 %) + Évaluation continue et régulière (TD et TP) (40 %).

7. Contenu

Ce cours de Méthodes d'échantillonnage présente en huit chapitres. En outre, chaque chapitre est traité à travers des séquences pédagogiques permettant l'assimilation des concepts prévus, cette assimilation est consolidée par des activités d'apprentissages où ces notions sont mises en œuvre en renforçant le cours. Le contenu de la matière est représenté ci-dessous :

Table des matières

- 1 – Introduction et rappels statistiques
- 2 – Théories de l'échantillonnage
 - 2.1 – Distribution de l'échantillonnage
 - 2.2 – Distribution et relation d'échantillonnage des caractéristiques statistiques
- 3 – Estimation des paramètres d'un lot
 - 3.1 – Estimation de la moyenne
 - 3.2 – Estimation de la variance et de l'écart type
 - 3.3 – Estimation de l'intervalle de confiance
- 4 – Détermination de la taille d'un échantillon
 - 4.1 – Précision pour une moyenne
 - 4.2 – Précision pour un lot
 - 4.3 – Les risques d'erreur
- 5 – Les méthodes d'échantillonnage
 - 5.1 – Les problèmes de l'échantillonnage, définitions et notions
 - 5.2 – Les erreurs lors de l'échantillonnage
 - 5.3 – L'effectif optimal des échantillons
 - 5.4 – Les méthodes d'échantillonnage
 - 5.4.1 – L'échantillonnage aléatoire simple
 - 5.4.2 – L'échantillonnage systématique
 - 5.4.3 – L'échantillonnage par grappe
 - 5.4.4 – L'échantillonnage par stratification
 - 5.4.5 – L'échantillonnage à deux degrés et à deux phases
- 6 – Elaboration d'un plan d'échantillonnage
- 7 – Comparaison des différentes méthodes d'échantillonnage
- 8 – Notions sur les enquêtes
- 9 – Références bibliographiques

CONTENU DU COURS:

Méthodes d'échantillonnage



1. INTRODUCTION ET RAPPELS STATISTIQUES

La statistique est l'étude de la collecte de données, leur analyse, leur traitement, l'interprétation des résultats et leur présentation afin de rendre les données compréhensibles par tous. C'est à la fois une science, une méthode et un ensemble de techniques.

L'analyse des données est utilisée pour d'écrire les phénomènes étudiés, faire des prévisions et prendre des décisions à leur sujet. En cela, la statistique est un outil essentiel pour la compréhension et la gestion des phénomènes complexes.

Les données étudiées peuvent être de toute nature, ce qui rend la statistique utile dans tous les champs disciplinaires et explique pourquoi elle est enseignée dans toutes les filières universitaires, de l'agronomie, l'économie à la biologie en passant par la psychologie et bien sûr les sciences de l'ingénieur. La statistique consiste à :

- ✓ Recueillir des données ;
- ✓ Présenter et résumer ces données ;
- ✓ Tirer des conclusions sur la population étudiée et d'aider à la prise de décision ;
- ✓ En présence de données dépendant du temps, nous essayons de faire de la prévision.

On appelle **population** la totalité des unités de n'importe quel genre prises en considération par le statisticien. Elle peut être finie ou infinie. Un **échantillon** est un sous-ensemble de la population étudiée.

Pour traiter un échantillon ou une population, le statisticien décrit habituellement ces ensembles à l'aide de mesures telles que le nombre d'unités, la moyenne, l'écart-type et la fréquence.

- ✓ Les mesures que l'on utilise pour décrire une population sont des paramètres. Un paramètre est une caractéristique de la population ;
- ✓ Les mesures que l'on utilise pour décrire un échantillon sont appelées des statistiques. Une statistique est une caractéristique de l'échantillon.

1.1. Quelques définitions

1.1.1. Définir les modalités de l'échantillonnage

L'échantillonnage est fondamental et résulte de l'impossibilité de collecter des données sur tous les éléments d'une population ou d'une surface, souvent pour des raisons pratiques, techniques ou économiques.

L'échantillonnage permet alors d'étudier le tout par le biais des statistiques. Il est pourtant l'un des aspects les plus négligés de la biostatistique, c'est ce qu'on peut constater aussi dans les espaces naturels.

La partie de la population que l'on va examiner s'appelle l'échantillon. Définir **les modalités de l'échantillonnage** consiste à définir la localisation, le nombre et la taille des échantillons de la population statistique.

Le recours à un spécialiste est souvent nécessaire pour aider le gestionnaire à définir la localisation et la densité des échantillons, la périodicité et la durée de l'échantillonnage. Il faut également se soumettre aux contraintes budgétaires et trouver un échantillon optimal pour baisser le coût des collectes et augmenter la précision des résultats.

1.1.2. Définition et objectif de l'échantillonnage

L'échantillonnage consiste essentiellement à tirer des informations d'une fraction d'un grand groupe ou d'une population, de façon à en tirer des conclusions au sujet de l'ensemble de la population. Son objet est donc de fournir un échantillon qui représentera la population et reproduira aussi fidèlement que possible les principales caractéristiques de la population étudiée.

L'objectif de l'échantillonnage influencera le type de plan d'échantillonnage, la nature des paramètres demandés et l'interprétation du résultat des essais.

Il est nécessaire, en tout premier lieu, de tenir compte du contexte et d'établir le but poursuivi lors du prélèvement des échantillons :

- Appui au travail du personnel d'inspection;
- Surveillance ou étude de profil;
- Investigation lors d'une toxi-infection alimentaire ou d'une plainte d'un consommateur;
- Recherche de pathogènes dans l'environnement;

- Poursuites, saisies, confiscations;
- Vérification de l'efficacité des méthodes de nettoyage et d'assainissement;
- Suivi de prélèvement relatif aux résultats analytiques antérieurs non conformes.

Il existe deux types d'échantillonnage : échantillonnage **exhaustif** (sans remise) et échantillonnage **non exhaustif** (avec remise).

Combien d'échantillons de **n** éléments peuvent être isolés d'une population de **N** éléments?

On distingue deux cas de tirage :

- Tirage exhaustif (sans remise) : nombre d'échantillons est C_N^n
- Tirage non exhaustif (avec remise) : nombre d'échantillons est N^n

1.1.3. Élément ou unité d'échantillonnage

Portion ou contenant individuel de produit prélevé au hasard dans un lot. Une unité d'échantillonnage peut correspondre à un échantillon.

Une des notions élémentaires de la théorie de l'échantillonnage, qu'il faut définir clairement pour construire le cadre d'échantillonnage, est l'unité d'échantillonnage, à savoir l'unité minimale d'observation à partir de laquelle on obtiendra de l'information sur les variables utiles.

Par convention, on emploiera en statistique le symbole "**N**" pour désigner le nombre d'unités d'échantillonnage qui intègrent le total et "**n**" pour désigner le nombre des individus dans l'échantillon considéré. On pourrait définir d'autres unités d'échantillonnage en fonction des objectifs de chaque étude.

Prenons tous les échantillons possibles de taille **n** tirés d'une population donnée de taille **N**. Pour chaque échantillon, on peut calculer une statistique (moyenne, écart-type, variance, etc.) qui variera avec l'échantillon. Pour tous les échantillons, on obtient alors une distribution de la statistique que l'on nomme **la distribution d'échantillonnage**. Pour la validité des résultats, il est important que les échantillons soient représentatifs de la population concernée.

1.1.4. Echantillon

Quantité de produit prélevé et représentatif d'un lot qui sera soumis à des essais en laboratoire. Un échantillon peut consister en une ou plusieurs unités d'échantillonnage.

L'échantillon est une collection d'éléments prélevés dans la population statistique (partie de la population que l'on va examiner) selon un processus aléatoire ou une méthode dite à choix raisonné. C'est le fragment d'un ensemble pour juger de cet ensemble.

L'échantillon doit être **représentatif** de cette dernière, c'est-à-dire, qu'il doit refléter fidèlement sa composition et sa complexité et fournir une estimation précise et non biaisée des paramètres mesurés sur les objets dans une aire donnée, à un moment donné.

Les résultats d'une étude sont d'autant plus fiables que le nombre de données à traiter est important. Ce dernier dépend de l'intensité des prélèvements, donc du nombre d'échantillons.

Une ou plusieurs unités d'échantillonnage prélevées aléatoirement dans une population et destinées à fournir des informations sur cette population.

Selon les principes de l'échantillonnage: Pas de choix d'un plan d'échantillonnage a priori Car son efficacité dépend du taux d'unités non conformes dans le lot de production contrôlé.

Pour un plan d'échantillonnage quel qu'il soit, il existe le risque d'accepter un lot non satisfaisant (contenant des unités non conformes), ce risque étant d'autant plus élevé que la proportion d'unités non conformes (unités contaminées à des taux inacceptables) dans le lot est faible.

1.1.5. Cadre d'échantillonnage

Dans le domaine agroalimentaire, les populations sont des animaux, des plantes ou des produits issus de ces derniers. Ces populations sont, dans la plupart des cas, subdivisées en troupeaux ou en lots. Des volumes de liquide peuvent aussi représenter un lot. Elles doivent être définies en fonction du temps et de l'espace. Cette définition s'appelle cadre d'échantillonnage.

Il est essentiel de définir le cadre d'échantillonnage toutes les fois que l'on désire décrire une population à l'aide d'un échantillon. L'échantillon décrit la population dont il est issu pour une période et un espace précis.

1.1.6. Population statistique

Une population statistique est une collection d'éléments, possédant au moins une caractéristique commune, permettant de la définir, de laquelle on extrait un échantillon représentatif et sur laquelle portent les conclusions statistiques.

1.1.7. Population cible

Un ensemble plus vaste d'où est extraite la population statistique. C'est généralement la population biologique qui comprend outre la population disponible, la population inaccessible et non vulnérable.

1.1.8. Variable

Caractéristique mesurée ou observée sur chacun des éléments de l'échantillon (variables propres: nombre d'individus, masse, etc.) ou sur des entités prédéfinies qui se rattachent aux unités d'échantillonnage (variables associées: température ambiante).

1.1.9. Paramètre

Caractéristique quantitative qui permet une représentation condensée de l'information contenue dans un ou plusieurs ensembles de données. Il s'agit d'un concept mathématique. Exemple: moyenne, écart-type, diversité, etc.

1.1.10. Lot

Partie d'une même livraison ou d'une même fabrication présentant les mêmes caractéristiques.

Les denrées alimentaires préemballées comportent obligatoirement la référence du lot de fabrication ou de conditionnement, un système harmonisé dans toute l'Union européenne. Cette référence du lot permet au consommateur une traçabilité du produit en cas de litige ou de danger pour la santé. La date de durabilité minimale ou la date limite de consommation peut faire office de mention permettant d'identifier le lot.

La directive **89/396/CEE** s'applique à toutes les **denrées alimentaires préemballées**.

Le terme "lot" désigne un ensemble d'unités de vente d'une denrée alimentaire produite, fabriquée ou conditionnée dans des circonstances identiques.

Elle ne s'applique pas aux:

- produits agricoles vendus à des stations d'entreposage, de conditionnement ou d'emballage, à des organisations de producteurs, aux systèmes de préparation ou de transformation;
- produits non préemballés;
- produits conditionnés dans un emballage ou un récipient inférieur à 10 cm² dans la face la plus grande;
- doses individuelles de glaces alimentaires.

→Étiquetage du lot

Le "lot", qui figure sur l'emballage ou le récipient, est déterminé par le producteur, le fabricant ou le conditionneur de la denrée alimentaire en question ou par le premier vendeur établi à l'intérieur de l'Union européenne.

L'indication est précédée de la lettre "L", sauf dans le cas où elle se distinguerait clairement des autres indications d'étiquetage. Dans tous les cas, elle figure de manière à être facilement visible, clairement lisible et indélébile.

2. THÉORIES DE L'ÉCHANTILLONNAGE

C'est l'étude des liaisons existant entre une population et les échantillons de cette population.

2.1. Distribution de l'échantillonnage

La distribution d'échantillonnage est l'étude de la probabilité de l'échantillon en fonction de la distribution de la variable parente lorsque la taille de l'échantillon augmente.

Pour résoudre les problèmes d'estimation de paramètres inconnus, il faut tout d'abord étudier les distributions d'échantillonnage.

Considérons que tous les échantillons possibles de taille n extraits d'une population de taille N , de moyenne μ , de variance σ^2 , etc. Pour chaque échantillon, il est possible de calculer les paramètres statistiques \bar{x} , s , s^2 , etc. qui varient d'un échantillon à l'autre. Chaque paramètre possédera ainsi une distribution d'échantillonnage au même titre que la variable aléatoire X .

On utilise souvent:

	Population	Echantillon	Distribution d'échantillonnage des moyennes	Distribution d'échantillonnage des variances
Taille	N	n	/	/
Moyenne	μ	\bar{X}	$\mu_{\bar{X}}$	μ_{s^2}
Ecart-type	σ	s	$\sigma_{\bar{X}}$	σ_{s^2}
Variance	σ^2	S^2	$\sigma^2_{\bar{X}}$	$\sigma^2_{s^2}$
Proportion	π	p	/	/

2.2. Distribution et relation d'échantillonnage des caractéristiques statistiques

2.2.1. Distribution d'échantillonnage des moyennes

Soit une population de taille N , on désigne par μ et σ la moyenne et l'écart-type de cette population respectivement. On extrait de la population une série d'échantillons de taille n , chacun de ces échantillons a une moyenne, les différentes moyennes obtenues

constituent une distribution d'échantillonnage de moyenne \mathbf{X} , on désigne par μ_x et σ_x la moyenne et l'écart-type de la distribution d'échantillonnage de la moyenne.

On a : $\mu_{\bar{x}} = \mu$

- Si le tirage est exhaustif (sans remise) : $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$
- Si le tirage est non exhaustif (avec remise) : $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Remarques

- Si n est plus petit devant N , la distinction entre exhaustivité et non exhaustivité est sans objet car $N-n/n-1 \approx 1$;
- Si la taille des échantillons est assez grande ($n \geq 30$), la distribution d'échantillonnage de moyennes approche la distribution normale quelle que soit la distribution de la population ;
- Si la population est normalement distribuée, la distribution d'échantillonnage de moyenne est une loi normale quelle que soit la valeur n de la taille des échantillons.

Exemple d'application (exemple pour le but de confirmer les formules précédentes)

Nous avons une population constituée de 3 éléments.

$$P = (1, 2, 3)$$

- Calculer la moyenne et l'écart-type de cette population.

Réponses

$$\text{On sait que } \mu = \frac{\sum_{i=1}^n x_i}{N} = \frac{1+2+3}{3} = 2$$

$$\text{Et on sait aussi que } \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N} = \frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2$$

$$\text{Donc : } \sigma^2 = \frac{1+4+9}{3} - 4 = \frac{2}{3}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{2}{3}}$$

$$\text{Résultats finaux : } \mu = 2 \text{ et } \sigma = \sqrt{\frac{2}{3}}$$

On va effectuer des prélèvements des échantillons de taille ($n = 2$), on fait le prélèvement dans les deux cas (échantillonnage exhaustif et échantillonnage non exhaustif).

- Quel est le nombre des échantillons qui peuvent être prélevés à partir de cette population ?
- Effectuer le prélèvement de ces échantillons.
- Etablir une distribution d'échantillonnage des moyennes.
- Calculer la moyenne de la distribution d'échantillonnage des moyennes.
- Calculer l'écart-type de la distribution d'échantillonnage des moyennes.
- Peut-on nous confirmer les deux formules ci-dessous.

Réponses

Cas 1 (échantillonnage non exhaustif / échantillonnage avec remise):

- *Parce que le tirage est avec remise donc le nombre des échantillons possibles à être prélever est une liste n éléments pris parmi N éléments c.à.d. $N^n = 3^2 = 9$*

- *Les échantillons sont les suivants :*

$$- \epsilon = \left\{ \begin{array}{l} (1, 1)(1, 2)(1, 3) \\ (2, 1)(2, 2)(2, 3) \\ (3, 1)(3, 2)(3, 3) \end{array} \right\}$$

- *La distribution d'échantillonnage des moyennes :*

$$- \left\{ \begin{array}{l} (1)(1, 5)(2) \\ (1, 5)(2)(2, 5) \\ (2)(2, 5)(3) \end{array} \right\}$$

- *La moyenne de la distribution d'échantillonnage des moyennes :*

$$- \mu_{\bar{x}} = \frac{\sum_{i=1}^n x_i}{N} = \frac{18}{9} = 2 = \mu$$

- *L'écart-type de la distribution d'échantillonnage des moyennes*

$$\sigma_{\bar{x}}^2 = \frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2 = \frac{1}{3}$$

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{3}}$$

- **Confirmation**

$$\mu_{\bar{x}} = \mu = 2$$

Formule confirmée

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{\frac{2}{3}}}{\sqrt{2}} = \sqrt{\frac{1}{3}} \quad \text{Formule confirmée}$$

Cas 2 (échantillonnage exhaustif / échantillonnage sans remise):

- Parce que le tirage est avec remise donc le nombre des échantillons possibles à être prélevé est une liste n éléments pris parmi N éléments c.à.d. $C_N^n = 3$

- Les échantillons sont les suivants :

$$- \epsilon = \{(1, 2)(1, 2)(2, 3)\}$$

- La distribution d'échantillonnage des moyennes :

$$- \{(1, 5)(2)(2, 5)\}$$

- La moyenne de la distribution d'échantillonnage des moyennes :

$$- \mu_{\bar{x}} = \frac{\sum_{i=1}^n x_i}{N} = \frac{6}{3} = 2 = \mu$$

-

- L'écart-type de la distribution d'échantillonnage des moyennes

$$\sigma_{\bar{x}}^2 = \frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2 = \frac{1}{6}$$

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{6}}$$

- Confirmation

$$\mu_{\bar{x}} = \mu = 2 \quad \text{Formule confirmée}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{\sqrt{\frac{2}{3}}}{\sqrt{2}} \sqrt{\frac{1}{2}} = \sqrt{\frac{1}{6}} \quad \text{Formule confirmée}$$

2.2.2. Distribution d'échantillonnage des variances

Chaque échantillon de taille n de la population à une variance S^2 , ces variances sont des valeurs observées d'une même variable aléatoire.

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{On a :}$$

$$E(S)^2 = \frac{n-1}{n} \sigma^2$$

$$V(S)^2 = \frac{n-1}{n^3} [(n-1)\mu^4 - (n-3)\sigma^4]$$

2.2.3. Distribution d'échantillonnage des fréquences

La probabilité de la réalisation d'un événement est supposée égale à p . nous considérons les échantillons de taille n extraits d'une population de taille N . à chaque échantillon extrait correspond une fréquence f_n de réalisation de l'événement considéré. Nous avons deux cas :

Si l'échantillonnage est avec remise :

$$E(f_n) = p$$

$$V(f_n) = \frac{p(1-p)}{n}$$

Si l'échantillonnage est sans remise :

$$E(f_n) = p$$

$$V(f_n) = \frac{p(1-p)}{n} \frac{N-n}{N-1}$$

2.2.4. Distribution d'échantillonnage des différences des moyennes

On considère 2 populations P_1 et P_2 de moyennes μ_1 et μ_2 et de variances σ^2_1 et σ^2_2 . On s'intéresse à la différence $\mu_1 - \mu_2$.

Nous avons :

$$\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$$

$$\sigma^2_{\bar{x}_1 - \bar{x}_2} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}$$

3. ESTIMATION DES PARAMETRES D'UN LOT

3.1. Estimation de la moyenne

La moyenne c'est la somme des valeurs observées divisées sur leur effectif total.

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{N}$$

X_i : Valeur observée.

N : Effectif total.

3.2. Estimation de la variance et de l'écart type

La variance σ^2 s'agit de la moyenne des carrés des écarts à la moyenne.

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N} = \frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2$$

L'écart-type σ est le plus utilisé des paramètres de dispersion. L'écart-type est beaucoup plus représentatif mais pour l'utiliser dans les cartes de contrôle en production il faut alors plutôt prévoir un calcul automatique.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}}$$

Le coefficient de variation **CV** représente une sorte d'écart-type relatif pour comparer les dispersions indépendamment des valeurs de la variable. Il s'exprime souvent en pourcentage.

$$CV = \frac{\sigma}{\bar{x}}$$

Le coefficient de variation permet de comparer notamment la précision de différents dosages effectués avec le même appareil.

3.3. Estimation de l'intervalle de confiance

Nous avons vu que la moyenne d'un échantillon aléatoire permet d'estimer la vraie moyenne de la population. Nous voudrions estimer également la précision de cette moyenne, c'est-à-dire donner une marge d'erreur ou intervalle de confiance.

Nous pouvons utiliser les tables de la loi normale pour estimer ces intervalles de confiance. En général, nous adopterons l'intervalle de confiance à 95%. Nous pouvons donc écrire :

$$\mu = \bar{x} \pm 2\sigma(\bar{x})$$

Et plus explicitement :

Il y a 95 chances sur 100 que μ se situe entre

$$\bar{x} - 2\sigma(\bar{x}) \text{ et } \bar{x} + 2\sigma(\bar{x})$$

4. DÉTERMINATION DE LA TAILLE D'UN ÉCHANTILLON

Lorsqu'on est maître de l'échantillonnage dans un travail planifié par exemple, on peut souhaiter obtenir une estimation avec précision spécifiée pour un seuil de confiance donné. La loi des grands nombres indique qu'il suffit d'extraire un échantillon d'effectif suffisant pour atteindre sûrement une précision donnée. On distingue 2 types de précisions :

- Une précision absolue ;
- Une précision relative (**k %**).

4.1. Précision pour une moyenne

Si **k %** est précision relative souhaitée, pour une confiance $1 - \alpha$, il faut réaliser la condition suivante :

$$|x - \mu| \leq Z_{(1-\alpha/2)} \sigma(\bar{X}) \sqrt{1 - n/N} \leq k\mu$$

$$|x - \mu| \leq Z_{(1-\frac{\alpha}{2})} \sigma / \sqrt{n} \sqrt{1 - n/N} \leq k\mu$$

En élevant au carré les termes de la seconde inégalité :

$$Z^2_{(1-\alpha/2)} \sigma^2 (N - n) \leq k^2 \mu^2 n N$$

Si **n/N** est très faible, la population peut être considérée comme infinie, le produit $Z^2_{(1-\alpha/2)} \sigma^2$ est négligeable devant $K^2 \mu^2 N$. L'expression se simplifie :

$$n \geq Z^2_{(1-\alpha/2)} \sigma^2 / k^2 \mu^2$$

Le calcul ne nécessite pas de connaître l'effectif total de la population.

Très souvent, on connaît l'ordre de grandeur du coefficient de variation $CV = \sigma/\mu$, l'expression précédente devient :

$$n \geq Z^2_{(1-\alpha/2)} CV^2 / k^2$$

Si l'échantillonnage est exhaustif, le calcul de l'effectif minimal nécessite en toute rigueur de connaître l'effectif de la population. Comme dans la plupart des cas il est inconnu, on détermine seulement un ordre de grandeur en supposant l'échantillon non exhaustif, ordre de grandeur qui surestime l'effectif nécessaire.

Exemple d'application:

Nous avons une précision relative à 5 % d'une population, avec une confiance de 0.95 et un coefficient de variation de 30 %.

- ❖ Quelle doit être la taille minimale d'un échantillon pour estimer la moyenne de cette population ?

Réponse :

CV = 0,30

k = 0,05 (précision relative)

Z = 1,96 ≈ 2

n = 2²(0,3)²/(0,05)² = 144

n = 144 c'est la taille minimale de l'échantillon.

4.2. Précision pour une proportion

Une détermination identique peut être faite pour une proportion où l'on souhaite une précision relative de k %. Si l'effectif est supposé être assez grand, l'approximation normale est possible. L'erreur aléatoire doit être inférieure à k p.

$$Z_{(1-\alpha/2)} \sqrt{p(1-p)/n} \sqrt{1-n/N} \leq kp$$

En opérant comme une moyenne :

$$n \geq Z^2_{(1-\alpha/2)} p(1-p)N / [k^2 p^2 N + Z^2_{(1-\alpha/2)} p(1-p)N]$$

Si l'échantillonnage peut être assimilé à un échantillonnage non exhaustif, la racine carrée de $1 - n/N$ est négligeable, l'expression se simplifie :

$$n \geq \frac{Z^2 (1 - p)}{k^2 p}$$

5. LES MÉTHODES D'ÉCHANTILLONNAGE

5.1. Les problèmes de l'échantillonnage

Il est rarement possible d'étudier de manière exhaustive une population pour différentes raisons entre autre de coût. On cherche donc à en réunir une fraction : un échantillon et à partir de ses caractéristiques statistiques à estimer les paramètres de la population. Pour que l'extrapolation ou l'inférence soit possible, l'échantillon doit être représentatif de la population. Par définition un échantillon est dit **représentatif** lorsqu'on peut, à partir de lui, décrire la population origine, en connaître la moyenne mais aussi la diversité.

L'échantillonnage est **représentatif** lorsque tout individu ou unité de la population peut figurer dans l'échantillon avec probabilité connue. Les procédures utilisées pour constituer les échantillons sont soumises à des contraintes statistiques mais aussi techniques. L'établissement d'un plan d'échantillonnage est alors nécessaire.

La plupart des méthodes d'échantillonnage se fondent sur la loi des grands nombres. Selon cette loi, la fréquence d'un événement converge en probabilité vers la probabilité de l'événement, la moyenne observée d'une variable tend vers son espérance mathématique (sa vraie valeur) lorsque l'effectif de l'échantillon augmente. Pour que la fréquence soit voisine de la proportion, pour que la moyenne soit proche de celle de la population étudiée, il suffit donc de réunir un échantillon d'effectif suffisant. On obtiendra alors une estimation **convenable** des valeurs recherchées. Mais le travail et le coût augmentant avec l'effectif, il est obligatoire limiter l'effectif.

❖ Définitions et notations

Considérons une population finie de N unités distinctes U_i repérées par exemple par un numéro ($i = 1, 2, 3, \dots, N$). Le terme unité doit être pris au sens large. Sur chacune d'elles on relève une caractéristique ou une variable X , soit X_i la valeur prise par chacune d'elles. On s'intéresse par exemple à la moyenne des unités, la proportion d'un caractère, l'effectif total, toute grandeur qui, pour la population, sont des valeurs certaines mais inconnues. (nous avons convenu de les désigner comme des paramètres statistiques et de les noter par des lettres grecques : μ, σ , etc). L'objectif est d'en obtenir une estimation et, si possible, de connaître la précision.

Un échantillon est une suite d'unités U_1, U_2, \dots, U_n prenant les valeurs X_1, X_2, \dots, X_n . Selon les méthodes, ce sont des unités d'observation ou de groupes d'unités : des unités d'échantillonnage. Les caractéristiques statistiques : moyenne, variance, fréquences calculées sur l'échantillon sont des variables aléatoires notées au moyen des lettres latines : x, s^2, f .

5.2. Les erreurs lors de l'échantillonnage

L'erreur d'échantillonnage se produit lorsqu'on estime une caractéristique de la population en étudiant seulement une partie de la population au lieu de la population au complet. Il s'agit de la différence entre l'estimation calculée à partir d'une enquête-échantillon et la « vraie » valeur qui aurait été obtenue si un recensement auprès de la population entière avait été effectué dans les mêmes conditions. **L'erreur d'échantillonnage varie avec la taille de l'échantillon.**

En général,

- l'erreur d'échantillonnage diminue au fur et à mesure que la taille de l'échantillon augmente (mais pas proportionnellement) ;
- L'erreur d'échantillonnage dépend de la taille de la population étudiée ;
- L'erreur d'échantillonnage dépend de la variabilité de la caractéristique de la population qu'on étudie ;
- On peut tenir compte de l'erreur d'échantillonnage et en réduire la portée grâce à un plan d'échantillonnage approprié ;
- L'erreur d'échantillonnage peut être mesurée et contrôlée dans les enquêtes qui utilisent un échantillon probabiliste.

Il existe des méthodes qui évaluent l'erreur d'échantillonnage pour les enquêtes portant sur un échantillon aléatoire. **La variance d'échantillonnage** est la mesure la plus couramment utilisée dans la quantification de l'erreur d'échantillonnage, et comme les autres méthodes, elle découle directement des méthodes d'échantillonnage et d'estimation utilisées dans l'enquête.

Il faut prendre en considération les risques d'erreurs :

1) Risques du client/consommateur

- ✓ Risque d'accepter un lot de mauvaise qualité ;
- ✓ Risque de considérer sous maîtrise un processus dérégulé ;

- ✓ En lien avec le Niveau de Qualité Limite.
- 2) Risques du fournisseur
- ✓ Risque de rejeter un lot de bonne qualité ;
 - ✓ Risque de considérer déréglé un processus maîtrisé ;
 - ✓ En lien avec le Niveau de Qualité Acceptable.

5.3. L'effectif optimal des échantillons

Préalablement à tout échantillonnage la question se pose de déterminer l'effectif minimal de l'échantillon conciliant différentes contraintes pratiques et permettant les estimations suffisamment précises souhaitées. L'effectif minimal diffère selon la nature du paramètre à estimer. L'estimation d'une précision requiert un effectif beaucoup plus élevé qu'une moyenne. Il appartient à l'expérimentateur de préciser l'erreur tolérée.

Dans un échantillonnage aléatoire, après avoir défini les unités d'échantillonnage, l'opérateur doit s'affranchir de toute réflexion pour le choix de celles à extraire et donc utiliser une méthode objective en recourant au hasard. Autrefois on utilisait des tables de nombres aléatoires : aujourd'hui on recourt le plus souvent aux ordinateurs ou aux calculettes capables de générer une suite de nombre aléatoires de 0 à 0,999. Pour générer une suite de nombres aléatoires x compris entre 1 et 125 par exemple, on utilise la formule suivante :

$$X = \text{INT} ((\text{RND} \times 125) + 1)$$

RND représente le nombre au hasard et **INT** indique que l'on retient la partie entière. sans la présence de l'unité, on n'obtiendrait que des nombres compris entre 0 et 124. Un programme informatique simple permet d'éliminer les unités désignées plusieurs fois.

L'extraction d'un échantillon d'effectif élevé exige un travail important, la méthode d'échantillonnage aléatoire est rarement utilisée comme telle mais la procédure aléatoire intervient dans d'autres méthodes.

L'échantillonnage idéal est un échantillonnage à 100%. Ce type d'échantillonnage est impossible pour des considérations économiques surtout lorsque 'il s'agit d'éléments indivisibles que l'analyse conduit à détruire (exemple : conserve). Toutefois, il peut être applicable dans le cas des produits en vrac ou une partie aliquote peut être prélevée sans nuire au reste du lot (exemple, sachet de semoule).

✓ **Echantillonnage satisfaisant**

Il est réalisé en prélevant un nombre d'échantillons égal à la V du nombre de divisions élémentaires, du produit à analyser. Exemple pour un lot de 10 000 boîtes de conserves, le nombre d'échantillons est égal à 100. Toutefois, ce nombre demeure toujours excessif.

Quel doit être la taille de votre échantillon ?

Il faut prendre en considération:

- La diversité de la population (plus la population est hétérogène plus la taille de l'échantillon doit être importante) ;
- Le temps dont on dispose ;
- Les coûts engendrés ;
- Les contraintes opérationnelles ;
- Le degré de précision.

5.4. Les méthodes d'échantillonnage

L'échantillonnage peut se faire avec ou sans remise et une population peut être considérée comme finie ou infinie. Une population finie dans laquelle on procède à un échantillonnage avec remise peut être théoriquement considérée comme infinie.

Dans la pratique, il en va de même pour des populations finies mais de grandes tailles. Pour chaque distribution d'échantillonnage, on peut calculer une moyenne, un écart type, une variance, etc.

Pour mettre en place des systèmes efficaces de contrôle des aliments, les organismes de réglementation et les laboratoires des États Membres doivent avoir la capacité d'appliquer des programmes d'échantillonnage bien conçus et des méthodes d'analyse éprouvées. L'AIEA concentre ses efforts sur l'élaboration de méthodes d'échantillonnage et d'analyse à transmettre à leurs laboratoires de réglementation et de recherche.

Afin de garantir la sécurité sanitaire et la qualité des denrées alimentaires, les États Membres ont besoin de systèmes de contrôle efficaces, d'organismes de réglementation et de laboratoires nationaux. Ceux-ci doivent disposer d'une infrastructure appropriée et de personnel bien formé, et avoir la capacité d'appliquer des programmes d'échantillonnage bien conçus et des méthodes d'analyse éprouvées.

En collaboration avec la FAO, l'AIEA aide les États Membres à développer leurs capacités en matière d'application de ce genre de programmes et de méthodes afin que les pratiques de sécurité alimentaire et environnementale mises en œuvre soutiennent l'intensification de la production végétale et la préservation des ressources naturelles.

→ Établissement de normes en vue de l'analyse en laboratoire et de l'échantillonnage

Les États Membres doivent disposer de méthodes d'analyse appropriées et validées, qui peuvent être appliquées par leurs laboratoires pour soutenir les systèmes de traçabilité des aliments, faciliter les tests d'authentification de ces derniers afin de lutter contre la fraude, et détecter et quantifier les contaminants alimentaires et environnementaux, notamment les résidus de médicaments vétérinaires ou de pesticides, ainsi que les toxines naturelles telles que les mycotoxines.

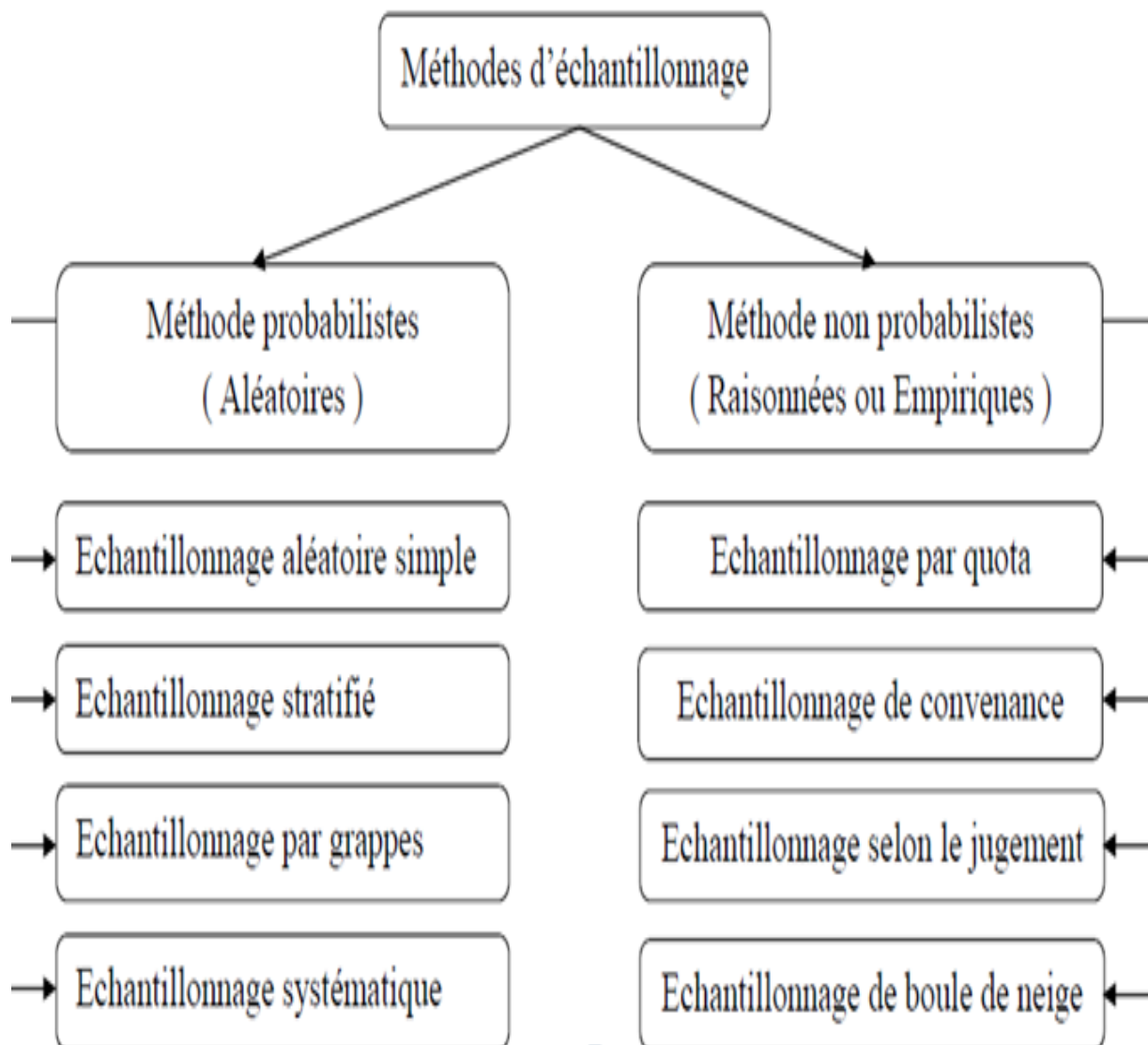
Afin que les résultats des analyses en laboratoire soient valables, il faut des plans et des procédures d'échantillonnage appropriés et validés pour l'utilisation prévue. Dans le domaine du contrôle réglementaire, par exemple, il est important de pouvoir estimer de manière précise les niveaux réels de contaminants présents dans les denrées alimentaires pour permettre de décider en connaissance de cause si elles peuvent ou non être consommées. Cela n'est possible qu'avec la collecte d'échantillons véritablement représentatifs.

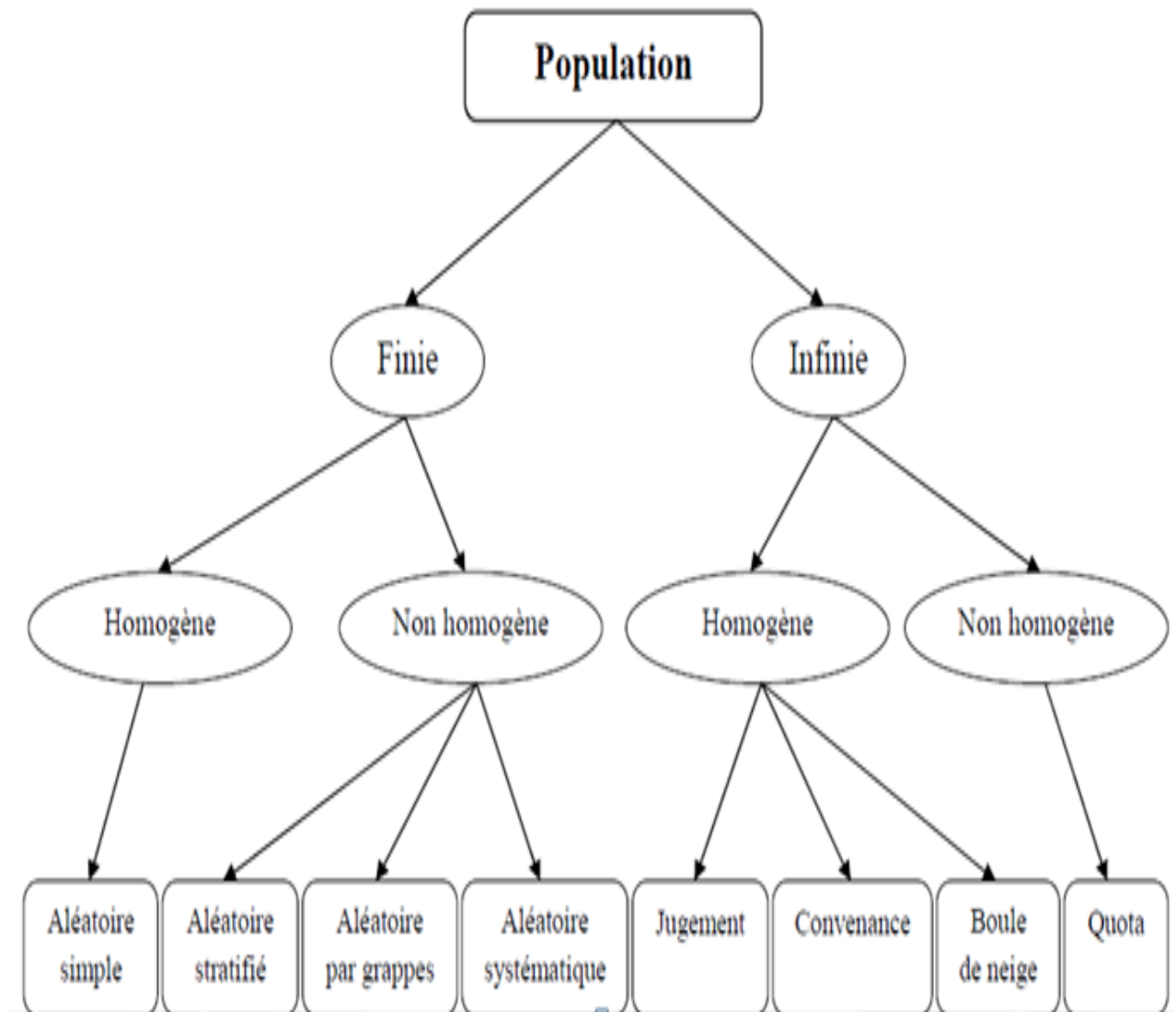
La Commission du Codex Alimentarius, créée en 1963 par la FAO et l'Organisation mondiale de la Santé pour mettre au point des normes alimentaires internationales harmonisées, élabore des normes et des lignes directrices, d'une part sur l'application de méthodes d'analyse et de plans d'échantillonnage pour lutter contre la contamination par les résidus de pesticides et de médicaments vétérinaires, et d'autre part sur d'autres questions ayant trait à la sécurité sanitaire, à l'authenticité et à la traçabilité des aliments. Le Système d'information sur les contaminants alimentaires et les résidus de la Division mixte FAO/AIEA des techniques nucléaires dans l'alimentation et l'agriculture comprend notamment des bases de données sur les méthodes d'analyse qui aident les États Membres à respecter les directives Codex.

L'AIEA élabore, adapte et valide des méthodes d'analyse afin qu'elles puissent être appliquées par les États Membres dans les laboratoires de réglementation et de recherche. Les techniques utilisées pour déterminer les contaminants présents dans les

aliments et vérifier l'authenticité de ces derniers comprennent la mesure des isotopes stables, l'établissement d'empreintes chimiques et la détermination du profil des éléments traces. Afin de permettre le contrôle de l'authenticité des aliments sur place, l'AIEA élabore des méthodes d'analyse recourant à des instruments de spectroscopie à main ou portatifs pouvant être utilisés sur le terrain.

Des méthodes multi-résidus et multi-contaminants sont utilisées autant que possible pour analyser les contaminants alimentaires, car elles sont économiques et satisfont aux normes de performance requises pour pouvoir être utilisées comme méthodes de réglementation visant à faciliter le commerce international. Les composés marqués par des isotopes stables sont utilisés comme normes internes pour promouvoir l'exactitude et améliorer la précision des méthodes utilisées en vue du respect des normes internationales.





5.4.1. Méthodes probabilistes (Aléatoires)

L'échantillonnage probabiliste repose sur un choix d'unités dans la population fait au hasard, ce n'est pas l'enquêteur qui choisit les unités, c'est la méthode utilisée pour la sélection qui le fait. Une des caractéristiques de cette méthode est que chaque unité de la population a une probabilité mesurable d'être choisie.

L'avantage de la méthode d'échantillonnage probabiliste est qu'elle permet de généraliser les résultats de l'échantillon à l'ensemble de la population en s'appuyant sur une théorie statistique reconnue.

Son seul inconvénient est qu'il faut posséder une liste de toutes les unités formant la population avant de procéder à la sélection de l'échantillon.

Voici les quatre types d'échantillonnage probabiliste que l'on peut effectuer :

5.4.1.1. Échantillonnage aléatoire simple

Un échantillon aléatoire simple est un échantillon sélectionné de manière à ce que chaque échantillon possible de taille " n " ait la même probabilité d'être sélectionné. L'échantillonnage aléatoire simple est la méthode d'échantillonnage la plus facile à appliquer et la plus couramment utilisée.

L'avantage de cette technique tient au fait qu'elle n'exige pas de données additionnelles dans la base de sondage (comme des régions géographiques) autres que la liste complète des membres de la population observée et l'information pour les contacter. Également, puisque l'échantillonnage aléatoire simple est une méthode simple et que la théorie qui la sous-tend est bien établie, il existe des formules-types pour déterminer la taille de l'échantillon, les estimations, etc., et ces formules sont faciles à utiliser.

D'un autre côté, cette technique ne fait aucunement appel aux données auxiliaires contenues dans la base de sondage (par exemple : le nombre d'employés de chaque entreprise) qui pourraient rendre le plan d'échantillonnage plus efficace. En outre, même s'il est facile d'appliquer l'échantillonnage aléatoire simple à de petites populations, le faire peut être coûteux et irréalisable pour de grandes populations, parce qu'il faut en identifier et en étiqueter toutes les unités avant l'échantillonnage. Son application peut également être coûteuse s'il faut effectuer des interviews sur place, puisqu'il est possible que l'échantillon soit géographiquement distribué dans toute la population. On prélève dans la population des individus au hasard, tous les individus ont la même probabilité d'être prélevés, et ils le sont indépendamment les uns des autres.

Pour prélever un échantillon aléatoire simple il faut:

- Constituer une base de sondage
- Numérotter les éléments de la population de **1** à **N**
- Procéder à sélection des éléments pour constituer un échantillon à l'aide de la **Table des nombres aléatoires**.

On utilise la table des nombres aléatoires en suivant 3 étapes:

1. On choisit aléatoirement un nombre dans la table compris entre **1** et **N** ;
2. On choisit notre chemin de lecture ;
3. On suit notre chemin, en choisissant seulement les nombres existants entre **1** et **N**.

Exemple:

On a une population de 850 individus et on veut choisir un échantillon de taille de 6 ($n = 6$).

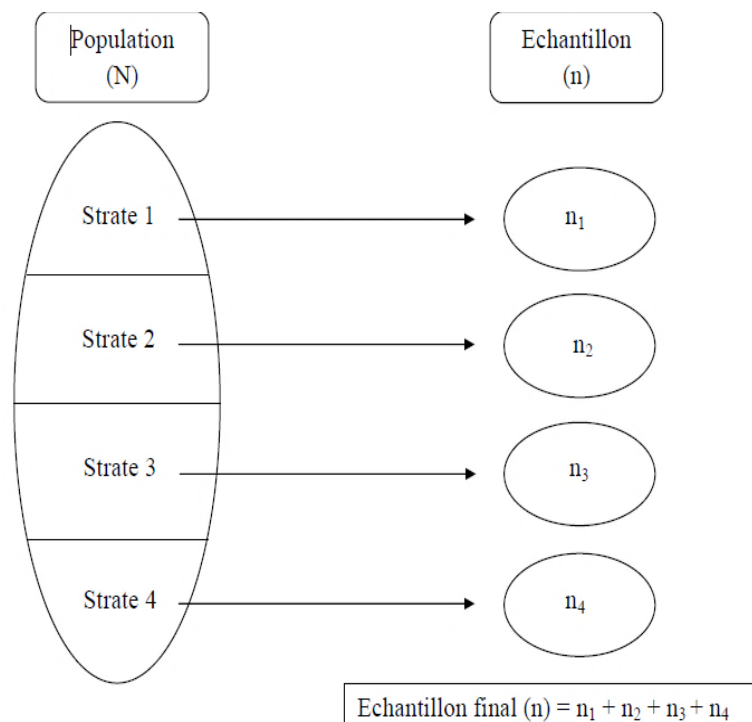
25760	20787	79572	18475
50537	55777	25547	24527
99795	18007	42862	77657
83753	71946	90985	95841
3200	56493	97953	29481

L'échantillon choisi est : (777; 7; 493; 572; 547; 475).

5.4.1.2. Echantillonnage aléatoire stratifié

On suppose que la population soit stratifiée, constituée de sous-populations homogènes, les strates. (Exemple : stratification par tranche d'âge). Dans chaque strate, on fait un échantillonnage aléatoire simple, de taille proportionnelle à la taille de strate dans la population (échantillon représentatif). Les individus de la population n'ont pas tous la même probabilité d'être tirés.

Le chercheur divise la population en sous-groupes distincts et homogènes (strates) à partir desquels il sélectionnera un échantillon aléatoire simple.



Exemple:

Supposons que 60% des laits sont poudre et 40% lait stérilisé, pour former un échantillon de 120 récipients en respectant ces strates, on devrait choisir au hasard $60\% \times 120 = 72$ lait en poudre et $40\% \times 120 = 48$ lait stérilisé.

a) Répartition proportionnelle**Exemple :**

Dans une population de 10000 entreprises réparties en 5000 petites entreprises, 3000 moyennes entreprises et 2000 grandes entreprises. On souhaite avoir un échantillon de 500 entreprises.

❖ Comment on peut répartir cet échantillon?

Strates	Effectif de la strate N	Effectif de l'échantillon t	Effectif de l'échantillon w
Petites entreprises	5000	$5000 \times 0.05 = 250$	$500 \times 0.5 = 250$
Moyennes entreprises	3000	$3000 \times 0.05 = 150$	$500 \times 0.3 = 150$
Grandes entreprise	2000	$2000 \times 0.05 = 100$	$500 \times 0.2 = 100$
Total	10000	$n = 500$	$n = 500$

❖ **Avantages et inconvénients de la méthode**

L'échantillonnage stratifié a l'avantage d'assurer une bonne représentation des différentes strates de la population dans l'échantillon. Il permet aussi d'obtenir des estimations pour chacune des strates de la population. Toutefois, pour utiliser cette méthode il faut avoir des renseignements sur la répartition des strates dans la population.

5.4.1.3. Echantillonnage aléatoire systématique

Dans certaines situations, spécialement lorsque les populations sont importantes, il est coûteux (en temps) de sélectionner un échantillon aléatoire simple en trouvant tout d'abord un nombre aléatoire et ensuite en cherchant dans la liste de la population l'élément correspondant. Une alternative de l'échantillonnage aléatoire simple est **l'échantillonnage systématique**.

Dans cette méthode on utilise la table des nombres aléatoires une seule fois pour déterminer h (le point d'entrée), les autres éléments qui constituent l'échantillon seront

déterminés à l'aide ce qu'on appelle le pas (**k**)($k=N/n$), l'échantillon sera prélevé comme suit: $k= (h, h+k, h+2k, h+3k, \dots)$ avec $1 < h < k$.

Exemple:

On veut sélectionner un échantillon de taille 3 au sein d'une population de taille N de 180.

$k = N/n = 180/3 = 60$, donc h doit compris entre 1 et 60 avec $1 < h < 60$.

25760	20787	79572	18475
50537	55777	25547	24527
99795	18007	42862	77657
83753	71946	90985	95841
3200	56493	97953	29481

L'échantillon choisi est : **(42; 102; 162)**.

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.

On utilise souvent cette méthode dans l'industrie, où l'on sélectionne une unité pour des essais à partir d'une chaîne de production afin de s'assurer que la machinerie et l'équipement sont d'une qualité uniforme. Un essayeur à l'intérieur d'une usine pourrait, par exemple, soumettre à un contrôle de la qualité chaque 20e produit sur une ligne de montage. L'essayeur pourrait choisir une origine au hasard entre les nombres 1

et 20. Cela déterminerait le premier produit à essayer; chaque 20e produit serait ensuite soumis à des essais.

Les intervieweurs peuvent utiliser cette technique d'échantillonnage lorsqu'ils interrogent des gens pour une enquête-échantillon. Le responsable d'une étude de marché pourrait sélectionner, par exemple, chaque 10e personne qui entrerait dans un magasin, après avoir sélectionné au hasard la première personne. Un enquêteur peut interviewer les occupants de chaque 5e maison d'une rue, après avoir sélectionné au hasard l'une des cinq premières maisons.

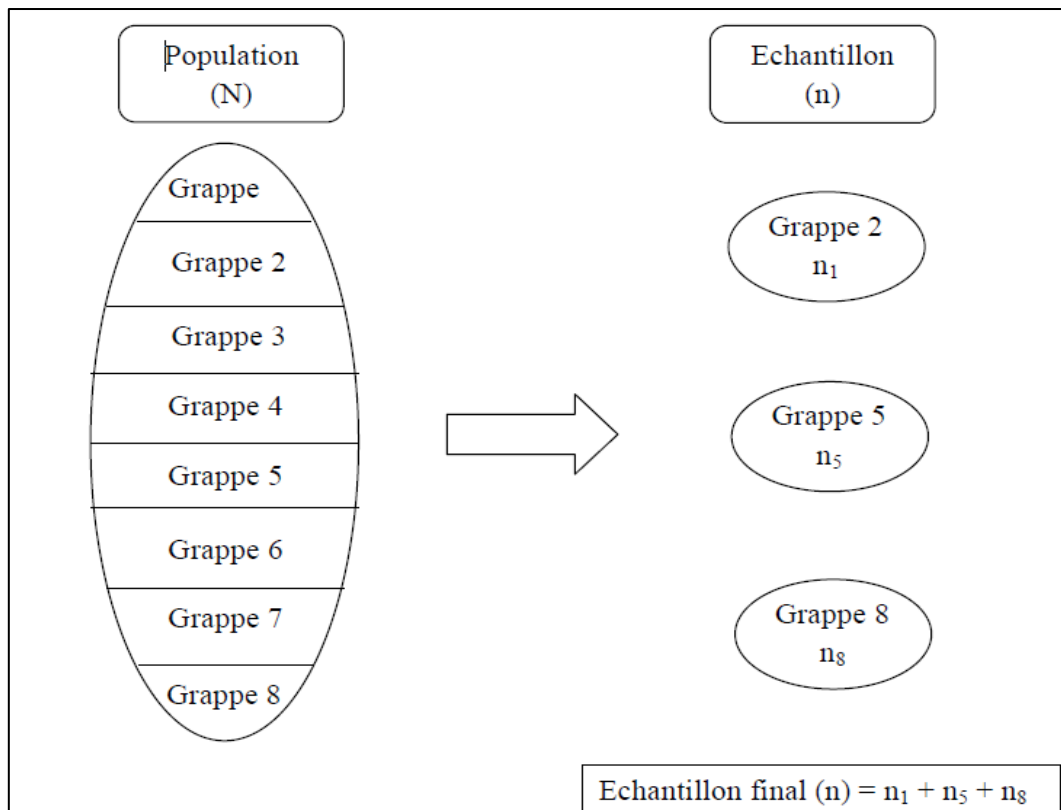
5.4.1.4. Echantillonnage aléatoire par grappe

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.

On tire au hasard des grappes ou familles d'individus, et on examine tous les individus de la grappe (exemple: on tire des immeubles puis on interroge tous les habitants). La méthode est d'autant meilleure que les grappes se ressemblent et que les individus d'une même grappe sont différents, contrairement aux strates.

Le chercheur divise la population en sous-groupes appelés « grappes ». Les grappes ont le même profil, la variance d'une grappe à l'autre étant faible. Il sélectionne par la suite un échantillon aléatoire de grappes et non pas un échantillon aléatoire à l'intérieur de chaque grappe.

**Exemple:**

Les récipients de jus sont répartis en 11 groupes, les groupes sont numérotés de 1 à 11. Supposons que l'on obtienne les nombres 2, 5, 7 et 10, tous les récipients de jus de ces 4 groupes feront partie de l'échantillon.

❖ **Avantages et inconvénients de la méthode**

L'avantage de cette méthode par rapport aux précédentes est qu'elle ne requiert pas au préalable la liste de la population, seule la liste des unités pour les grappes pignées est nécessaire. Un désavantage de ce type d'échantillonnage est qu'il produit des estimations habituellement moins précises que l'échantillonnage aléatoire simple parce que des unités appartenant à une même grappe ont tendance à présenter des caractéristiques semblables. Cette perte de précision peut être compensée par une augmentation de la taille de l'échantillon.

5.4.1.5. Echantillonnage aléatoire à plusieurs degrés

La méthode d'échantillonnage à plusieurs degrés ressemble à la méthode d'échantillonnage en grappes, sauf qu'il faut dans son cas prélever un échantillon à l'intérieur de chaque grappe sélectionnée, plutôt que d'inclure toutes les unités dans la

grappe. Ce type d'échantillonnage exige au moins deux degrés. On identifie et sélectionne au premier degré de grands groupes ou de grandes grappes. Ces grappes renferment plus d'unités de la population qu'il n'en faut pour l'échantillon final.

Pour obtenir un échantillon final, on prélève au second degré des unités de la population à partir des grappes sélectionnées (à l'aide de l'une des méthodes d'échantillonnage probabiliste possibles). Si l'on utilise plus de deux degrés, le processus de sélection d'unités de la population à l'intérieur des grappes se poursuit jusqu'à l'obtention d'un échantillon final.

5.4.1.6. Echantillonnage aléatoire à plusieurs phases

Un échantillonnage à plusieurs phases entraîne la collecte de données de base auprès d'un échantillon d'unités de grande taille et ensuite, pour un sous-échantillon de ces unités, la collecte de données plus détaillées. La forme la plus courante d'échantillonnage à plusieurs phases est l'échantillonnage à deux phases (ou l'échantillonnage double), mais il est également possible d'effectuer un échantillonnage à trois phases ou plus.

L'échantillonnage à plusieurs phases est assez différent de l'échantillonnage à plusieurs degrés, malgré les similarités entre eux sur le plan de leur appellation. Même si l'échantillonnage à plusieurs phases suppose aussi le prélèvement de deux échantillons ou plus, dans son cas, tous les échantillons sont tirés de la même base de sondage et les unités sont structurellement les mêmes à chaque phase. Comme dans le cas de l'échantillonnage à plusieurs degrés, plus l'on utilisera de phases, plus le plan d'échantillonnage et l'estimation deviendront complexes.

L'échantillonnage à plusieurs phases est utile lorsqu'il manque à l'intérieur de la base de sondage des données auxiliaires qui pourraient servir à stratifier la population ou à rejeter à la sélection une partie de la population.

5.4.2. Méthodes non probabilistes (Raisonnées ou empirique)

L'échantillonnage non probabiliste repose sur un choix arbitraire des unités, c'est l'enquêteur qui choisit les unités et non le hasard. En ce sens, il serait donc aventureux de généraliser les résultats obtenus pour l'échantillon à toute la population. La différence entre l'échantillonnage probabiliste et l'échantillonnage non probabiliste tient à une hypothèse de base au sujet de la nature de la population étudiée. Dans le cas de

l'échantillonnage probabiliste, chaque unité a une chance d'être sélectionnée. Dans celui de l'échantillonnage non probabiliste, on suppose que la distribution des caractéristiques à l'intérieur de la population est égale. C'est ce qui fait que le chercheur croit que n'importe quel échantillon serait représentatif et que les résultats, par conséquent, seront exacts.

Pour l'échantillonnage probabiliste, la randomisation est une caractéristique du processus de sélection, plutôt qu'une hypothèse au sujet de la structure de la population. Malgré cela, ces méthodes sont souvent utilisées dans certaines disciplines. En voici quelques-unes :

5.4.2.1. Echantillonnage par quotas

L'échantillonnage par quotas est l'une des formes les plus courantes d'échantillonnage non probabiliste. Il s'effectue jusqu'à ce qu'un nombre précis d'unités (de quotas) pour diverses sous-populations ait été sélectionné. Puisqu'il n'existe aucune règle qui régirait la façon dont il faudrait s'y prendre pour remplir ces quotas, l'échantillonnage par quotas est réellement un moyen de satisfaire aux objectifs en matière de taille d'échantillon pour certaines sous-populations.

Les quotas peuvent être fondés sur des proportions de la population. Si une population, par exemple, compte 100 hommes et 100 femmes et s'il faut en prélever un échantillon de 20 personnes pour qu'elles participent à un concours de dégustation de colas, il se peut que vous vouliez diviser l'échantillon en proportions égales entre les sexes, ce qui donnerait 10 hommes et 10 femmes. On peut penser que l'échantillonnage par quotas est préférable à d'autres formes d'échantillonnage non probabiliste (comme l'échantillonnage au jugé), parce qu'il impose l'inclusion dans l'échantillon de membres de différentes sous-populations.

L'échantillonnage par quotas est un peu similaire à l'échantillonnage stratifié parce que dans son cas également les unités semblables sont regroupées. Toutefois, il en diffère, cependant, sur le plan du mode de sélection. Dans le cas d'un échantillonnage probabiliste, on sélectionne les unités au hasard, tandis que dans celui d'un échantillonnage par quotas, on laisse habituellement à l'intervieweur le soin de déterminer qui sera échantillonné. Cela peut donner lieu à des biais de sélection. Les responsables d'études de marché utilisent donc souvent l'échantillonnage par quotas (pour des enquêtes ou des sondages téléphoniques, en particulier), plutôt que

l'échantillonnage stratifié, parce qu'il est relativement peu coûteux et facile à administrer et à la propriété souhaitable de respecter les proportions de la population. L'échantillonnage par quotas camoufle toutefois des biais pouvant être significatifs.

Comme dans le cas de toutes les autres méthodes d'échantillonnage non probabiliste, il faut supposer pour l'échantillonnage par quotas que les personnes sélectionnées sont semblables à celles qu'on ne sélectionne pas, afin de formuler des inférences au sujet de la population. Des hypothèses aussi audacieuses sont rarement valables.

La principale différence entre l'échantillonnage stratifié et l'échantillonnage par quotas tient au fait que le premier entraînerait la sélection des élèves à l'aide d'une méthode d'échantillonnage probabiliste comme l'échantillonnage aléatoire simple ou l'échantillonnage systématique. On n'utilise pas une telle technique dans le cas de l'échantillonnage par quotas. On pourrait sélectionner les 15 élèves en choisissant les 15 premiers élèves de 10^e année qui entreraient à l'école une journée donnée ou en choisissant 15 élèves dans les deux premières rangées d'une classe en particulier. N'oubliez pas que les élèves qui arrivent en retard ou qui s'assoient dans le fond de la classe peuvent avoir des opinions différentes de celles des élèves qui arrivent plus tôt à l'école ou qui s'assoient en avant lorsqu'ils entrent dans la classe.

Le fait que l'échantillonnage par quotas ne respecte pas l'exigence fondamentale du hasard est le principal argument militant contre son utilisation. Certaines unités peuvent n'avoir aucune chance d'être sélectionnées ou on risque de ne pas connaître leur chance de l'être. L'échantillon peut donc être biaisé.

Il est courant, mais il n'est pas nécessaire, que l'échantillonnage par quotas fasse appel à des procédures de sélection au hasard aux stades de départ, en grande partie de la même façon que le fait l'échantillonnage probabiliste. La première étape de l'échantillonnage à plusieurs degrés, par exemple, consisterait à sélectionner au hasard les régions géographiques. La différence se situe au niveau de la sélection des unités aux stades finals du processus.

Exemple d'application :

Une population de 15000 entreprises étudiées selon le chiffre d'affaire (X). Les résultats obtenus sont mentionnés dans le tableau ci-dessous.

X	Effectif (Ni)
0-50	5500
50-100	4500
100-200	1750
200-500	2000
500-1000	500
> 1000	750

Soit un taux de sondage de 0,10

1. Quelle méthode peut-on utiliser pour déterminer un échantillon ?
2. On suppose que l'institut dispose de 30 enquêteurs :
 - Calculer N_i/N
 - Déterminer la répartition au sein de l'échantillon
 - Déterminer la répartition des enquêteurs

Réponses :

1. La méthode utilisée c'est Echantillonnage par quotas car : pas de liste des éléments et il existe un taux de sondage

2.

X	Effectif (Ni)	Quotas population	Effectif échantillon	Répartition enquêteurs
0-50	5500	36.67%	550	11
50-100	4500	30.00%	450	09
100-200	1750	11.66%	175	03
200-500	2000	13.34%	200	04
500-1000	500	03.33%	50	01
> 1000	750	05.00%	75	02
Total	15000	100%	1500	30

5.4.2.2. Echantillonnage de convenance (de commodité)

Cas où les unités d'échantillonnage sont faciles à rejoindre, disponibles et généralement facile à convaincre. On appelle parfois l'échantillonnage de commodité l'échantillonnage à l'aveuglette ou accidentel. Cet échantillonnage n'est pas

normalement représentatif de la population cible, parce qu'on ne sélectionne des unités d'échantillonnage dans son cas que si on peut y avoir facilement et commodément accès.

Il arrive que Tout-le-monde utilise l'échantillonnage de commodité. Un critique gastronomique, par exemple, peut goûter plusieurs entrées ou plats principaux pour juger de la qualité et de la variété d'un menu. Les reporters des stations de télévision sont, en outre, souvent à la recherche de soi-disant « interviews de gens de la rue » pour déterminer comment la population perçoit un enjeu ou une question. Dans ces deux cas, on choisit l'échantillon au hasard, sans utiliser de méthode d'enquête particulière.

L'avantage évident de la méthode, c'est qu'elle est facile à utiliser, mais la présence de biais annule énormément ce dernier. Même si ses applications utiles sont limitées, la technique peut donner des résultats exacts lorsque la population est homogène. Un scientifique pourrait, par exemple, utiliser cette méthode pour déterminer si un produit alimentaire est contaminé. En supposant que l'eau du produit est bien mélangée, tout échantillon donnerait de l'information identique. Un scientifique pourrait en toute sécurité puiser de l'eau n'importe où dans le lac, sans se tracasser au sujet de la représentativité de son échantillon.

5.4.2.3. Echantillonnage selon le jugement

On utilise la méthode d'échantillonnage au jugé lorsqu'on prélève un échantillon en se fondant sur certains jugements au sujet de l'ensemble de la population. L'hypothèse qui sous-tend son utilisation est que l'enquêteur sélectionnera des unités qui seront caractéristiques de la population. La question cruciale dans ce cas est l'objectivité : Dans quelle mesure peut-on se fier à son jugement pour en arriver à un échantillon typique? L'échantillonnage au jugé est exposé aux préjugés du chercheur et est peut-être encore davantage biaisé que l'échantillonnage de commodité ou à l'aveuglette. Étant donné que l'échantillonnage au jugé reflète toutes les idées préconçues que risque d'avoir le chercheur, il peut y avoir introduction de biais importants si ces idées sont inexactes.

Les statisticiens utilisent souvent cette méthode dans le cadre d'études préparatoires comme des tests préalables de questionnaires et des discussions en groupe. Ils préfèrent également avoir recours à cette méthode à l'intérieur du cadre de laboratoires où le choix des sujets des expériences (comme des animaux, des êtres humains et des végétaux) reflète les croyances ou les convictions antérieures de

l'enquêteur au sujet de la population. La réduction du coût et du temps qu'exige l'acquisition de l'échantillon est l'un des avantages de l'échantillonnage au jugé.

5.4.2.4. Echantillonnage boule de neige

Dans le domaine de la recherche en statistique, l'échantillonnage boule de neige est une méthode d'échantillonnage non probabiliste dans laquelle les sujets de l'étude recrutent d'autres sujets parmi leur entourage. La population de l'échantillon s'étoffe à la manière d'une boule de neige qui roule. À mesure que la population de l'échantillon s'agrandit, les données atteignent un seuil suffisant pour être exploitées par les chercheurs.

L'échantillonnage boule de neige fonctionne avec un petit groupe d'informateurs chargés de recruter, via leurs communautés, d'autres participants correspondant aux critères et susceptibles de participer à une étude spécifique. Le terme « boule de neige » illustre l'analogie de cette méthode avec une boule de neige qui gagne de nouvelles couches à mesure qu'elle roule. Comme les membres de l'échantillon ne sont pas choisis dans une base de sondage, les échantillonnages boule de neige sont susceptibles de présenter de nombreux biais.

Ce type d'échantillonnage est utile dans le cas de la rareté des unités d'échantillonnage ou de l'absence d'un cadre d'échantillonnage valide. On demande à un répondant de nous référer à un autre qui présente les mêmes caractéristiques que les siennes, et ainsi de suite.

Dans, la méthode boule de neige, un ménage est identifié au départ et par incrémentation l'échantillon de ménage à enquêter se constitue. En fait, chaque individu (J) va permettre la sélection d'autres individus (V, M, Ab, I, R, ...). Ici, le choix du 1^{er} individu conditionne la composition de l'échantillon. C'est l'effet boule de neige.

6. ÉLABORATION D'UN PLAN D'ÉCHANTILLONNAGE

Il faut élaborer un plan d'échantillonnage précisant les conditions et la manière dont les unités seront prélevées.

Les caractéristiques statistiques de l'échantillon permettent d'estimer les paramètres statistiques de la population et leur précision. On applique la loi normale ou de *Student* selon l'effectif de l'échantillon réuni.

Le nombre d'échantillons à prélever est déterminé en fonction de la situation. Lorsqu'il est nécessaire de conclure sur un lot de produits, il faut s'entendre avec l'analyste du laboratoire afin de fixer le plan d'échantillonnage le plus approprié. Le plan d'échantillonnage est grandement influencé par le risque à la santé et à l'hétérogénéité du lot. Il arrive que ce plan d'échantillonnage soit déjà établi comme c'est le cas dans le cadre de la programmation analytique. À ce moment, il est possible de se référer à la documentation relative au programme.

6.1. Les éléments d'un plan d'échantillonnage

L'élaboration d'un plan d'échantillonnage suppose de déterminer les éléments suivants :

- ✓ L'objectif de l'étude et le choix de la variable à considérer ;
- ✓ L'effectif du ou des échantillons selon la précision souhaitée ;
- ✓ La base d'échantillonnage utilisée ;
- ✓ La méthode d'échantillonnage ;
- ✓ Le plan de prélèvement ;
- ✓ Le coût de réalisation.

Les éléments précédents ne sont pas totalement indépendants. La disposition préalable d'une base d'échantillonnage ou sa constitution plus ou moins facile peut orienter le choix d'une méthode. Inversement l'utilisation d'une méthode particulière peut imposer une certaine constitution de la base. On dispose parfois d'indicateur pour orienter le choix d'une méthode, c'est par exemple entre stratification et échantillonnage en grappes.

6.1.1. Définition des objectifs

L'objectif général est d'estimer les principaux paramètres de la population en appliquant successivement (et exceptionnellement) plusieurs méthodes d'échantillonnage. Les vraies valeurs sont, elles aussi, exceptionnellement connues pour permettre des comparaisons.

De plus, en réunissant une série d'échantillon pour chaque méthode on a cherché à comparer les méthodes et à montrer la normalité des distributions d'échantillonnage.

6.1.2. Choix d'un plan d'échantillonnage

Type de plan: de façon générale plan par attributs binomiale : n et c

1) A deux classes (mais pas exclusivement) : absence/présence, supérieur/inférieur à une limite :

- ✓ Agents très pathogènes en eux-mêmes ;
- ✓ Toxine.

2) A trois classes : (à trois classes) m , M , $c/n...$ et des précisions d'interprétation

- ✓ Germes témoins ;
- ✓ Toxine.

6.1.3. Précision et effectif des échantillons

Pour atteindre la précision souhaitée, à un niveau de confiance donné, il faut réunir un échantillon présentant un effectif minimal.

Rappelons qu'il faut deux manières d'exprimer la précision :

- La précision absolue ;
- La précision relative.

❖ Fixation de l'effectif pour l'étude comparative

Il apparaît difficile d'obtenir la précision relative de 5%, sans réunir un effectif très élevé du fait de la valeur élevée du coefficient de variation de la population. Il faut prendre en considération les facteurs suivants :

- **Choix d'un effectif divisible par de nombreux facteurs ;**
- **Nécessité d'un effectif assez élevé pour montrer la normalité des distributions d'échantillonnage ;**

- **Souci de la réalisation pratique, la masse des échantillons à manipuler doit être raisonnable.**

❖ **Nombre critique d'unités non conformes**

Choix dépend des niveaux de qualité acceptable et limite / risques fournisseur et client.

6.1.4. La base d'échantillonnage utilisée

Une base d'échantillonnage consiste en une liste, un fichier où les unités à échantillonner sont identifiées sans répétition et sans omission par un numéro, des coordonnées, etc.

❖ **Dispositions possibles**

Dans le lot, les sacs de maïs sont disposés en rangs et en lignes, la circulation n'est disponible qu'entre les rangs. Deux dispositions de la base d'échantillonnage sont possibles. En utilisant la ligne numéro 1 comme références, les sacs peuvent être identifiés au moyen de deux coordonnées : rang, ligne, comme l'indique la figure 1a.

On peut aussi envisager une numérotation continue comme l'indique la figure 1b.

1.1	2.1	3.1			461	47.1	48.1
1.2	2.2	3.2			46.2	47.2	48.2
1.3	2.3	3.3			46.3		
1.28	2.28	3.28					
1.29	2.29	3.29					
1.30	2.30	3.30			46.30	47.30	48.30

Figure 1a. Base d'échantillonnage : double identification (Ba).

1	31	61			1351	1381	1411
2	32	62				47.2	48.2
3	33	63					
28	58						
29	59						1439
30	60	90			1380	1410	1440

Figure 1b. Base d'échantillonnage : numérotation continue (Bb).

Quelle que soit l'étude, la répartition des coûts doit être :

- 10 à 15 % pour la conception ;
- 60 à 75 % pour la réalisation ;
- 10 à 15 % pour l'interprétation.

Attention

- ✚ À ne pas importer les critères d'autres pays ou union sans tout le travail d'analyse et de réflexion préalable pour s'assurer de la validité ;
- ✚ À ne pas passer automatiquement au plan à 2 classes ;
- ✚ À considérer que les critères étudiés peuvent évoluer dans leur sévérité dans le temps et dans l'espace.

L'utilisation de plans renforcés doit être envisagée d'abord sous l'angle de la fréquence.

Détection d'une non-conformité sur un produit ou un lot de produits ;

Mise en évidence d'un dysfonctionnement du procédé de fabrication ;

La taille de l'échantillonnage doit être déterminée au cas par cas, en fonction

- ✓ de la probabilité de détection du plan ;
- ✓ du niveau de prévalence estimé du danger considéré.

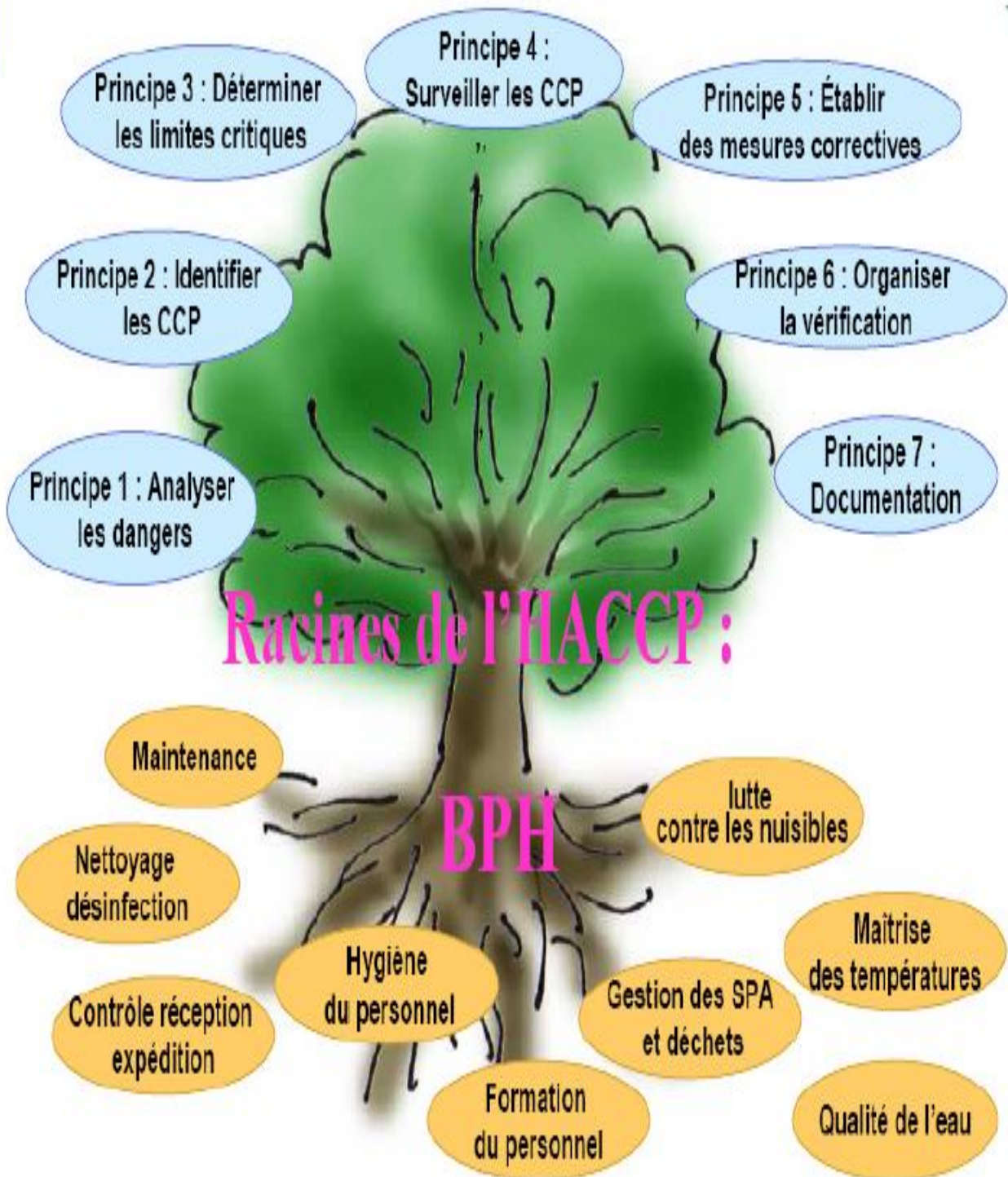
6.2. Limites des plans d'échantillonnage

Exemple : microbiologie des aliments

- ✓ Les analyses destinées à vérifier la conformité aux critères microbiologiques existants ne sont en aucun cas suffisantes pour garantir la sécurité sanitaire des aliments ;
- ✓ Dans tous les cas mettre en place une approche préventive (maîtrise de l'ensemble des paramètres susceptibles d'intervenir le long de la chaîne alimentaire) ;
- ✓ La gestion de la sécurité sanitaire des aliments a évolué ;
- ✓ Contrôle des produits finis ;
- ✓ Maîtrise des procédés ;
- ✓ fondée sur les dangers: BPH (A, F, V...) / HACCP

Les exemples suivants de plans d'échantillonnage à deux classes reposent sur :

- ✓ L'hypothèse favorable d'une distribution homogène du micro-organisme considéré dans le lot ;
- ✓ Cette hypothèse n'est généralement pas valable pour les aliments solides ; dans ce cas les probabilités ci-dessous d'accepter un lot contenant des unités non conformes sont plus élevées.



7. NOTIONS SUR LES ENQUETES

7.1. Définition de l'enquête

L'enquête est le recueil de données qui seront exprimées sous une forme statistique. L'enquête est une méthode de recueil de données primaires à partir d'un questionnaire administré à un échantillon issu d'une population cible.

Les enquêtes alimentaires sont des méthodes développées pour évaluer les apports alimentaires d'un individu, ou d'un groupe d'individus. L'évaluation des apports alimentaires est utilisée en épidémiologie et en pratique clinique, avec des objectifs un peu différents.

L'objectif principal de l'épidémiologie nutritionnelle est de mettre en relation les modes de consommations alimentaires et le risque de développer certaines pathologies. Les enquêtes permettent ainsi de cerner des nutriments, des aliments ou des profils de consommation plus ou moins bénéfiques ou néfastes à la santé. L'élaboration des apports nutritionnels conseillés pour la population, des doses toxiques maximales tolérables ou encore des guides de recommandations pour l'alimentation repose aussi sur les données des enquêtes alimentaires réalisées à grande échelle.

En clinique, l'évaluation des apports alimentaires fait, entre autre, partie de la prise en charge des maladies « liées à la nutrition » mises en évidence par l'épidémiologie nutritionnelle.

Le but des enquêtes alimentaires est de connaître qualitativement et quantitativement la consommation alimentaire. Elles peuvent être utilisées pour une population importante, pour un groupe restreint de sujets (enquêtes alimentaires collectives) ou pour un individu (enquêtes alimentaires individuelles). Les méthodes employées sont de deux types:

- ✓ évaluations de consommation globale (l'interprétation des statistiques de production en fonction du nombre et de la composition de la population considérée);
- ✓ études de consommation précise (méthodes des fréquences de consommation, relevé des achats et des pertes, techniques de la liste ou de l'inventaire, entretien-questionnaire, enquête pesée).

D'autres enquêtes recherchent les facteurs conscients et inconscients de la consommation alimentaire : enquêtes de motivations psychologiques, enquêtes socio-économiques, historiques (évolution du comportement alimentaire) ou géographiques (visant à décrire sa répartition). Elles utilisent toutes des questionnaires comportant des questions ouvertes ou fermées, des interviews individuelles ou de groupes, non directives ou semi-directives. Quel que soit leur but, les enquêtes alimentaires peuvent être rétrospectives ou prospectives. La validité de leurs résultats doit être évaluée en fonction du mode de recueil des informations et du but poursuivi. L'interprétation nécessite toujours beaucoup d'esprit critique et le mode d'expression des résultats est choisi pour rendre exactement compte de leur signification. L'intérêt de ces enquêtes réside, soit dans la photographie d'un comportement collectif ou individuel à un moment donné, soit dans l'étude de l'évolution d'un comportement ou de ses causes. Elles intéressent le planificateur, l'économiste, le producteur, le transformateur, le distributeur de l'aliment. En épidémiologie, elles constituent un élément d'un ensemble de données permettant d'évaluer l'éventualité, la nature et la force d'une corrélation de la consommation alimentaire ou de son évolution, soit avec les valeurs ou les évolutions de paramètres cliniques et biologiques, soit avec la prévalence et l'incidence d'une maladie. Technique de contrôle, technique de recherche, les enquêtes alimentaires de prix de revient souvent élevé, constituent le seul et irremplaçable moyen d'étude des conditions de nutrition, connaissance indispensable pour une politique logique de prévention.

7.2. Construction

Tout questionnaire d'enquête doit avoir une introduction présentant le but de la demande. Si le nombre de questions est très important, alors les questions pourront être divisées en grandes parties distinguées par des titres.

Les questions signalétiques (de renseignements) apparaissent toujours en fin de document. Les questions proposées peuvent être **ouvertes** (Elles demandent une réponse libre) ou **fermées** (elles demandent une réponse fixe unique ou multiple).

7.3. Méthodes de recueil des apports alimentaires

Le recueil des apports alimentaires peut être envisagé sous plusieurs angles. Nous décrirons, pour information, les méthodes d'enquêtes alimentaires réalisées au

niveau collectif en dernière partie de ce paragraphe, mais l'essentiel de cet exposé portera sur la description des méthodes recueillant les données au niveau de l'individu. Parmi les méthodes d'enquête alimentaire réalisées au niveau individuel, certaines recueillent les consommations sur des jours définis alors que d'autres s'attachent à recueillir des informations sur les consommations habituelles du sujet.

7.3.1. Recueil des apports sur des jours définis

Enregistrements alimentaires

L'enregistrement alimentaire a longtemps été considéré comme la méthode de référence parce qu'il permet d'apporter des informations précises sur les apports alimentaires. Dans ce type d'enquête, on demande au participant de noter sur un carnet le détail de ses consommations d'aliments et de boissons pendant une période déterminée.

Historiquement, l'enregistrement alimentaire était préconisé sur une période de 7 jours, de manière à couvrir les variations d'apports observées au cours d'une semaine. En pratique, il est fréquemment réalisé sur une période de 3 ou 4 jours pour éviter une perte de compliance des sujets liée à un enregistrement trop long. Pour faciliter l'organisation de l'enquête, le recueil se fait en général sur des jours consécutifs comprenant au moins un jour de week-end, mais certains protocoles imposent parfois qu'il soit réalisé sur des jours non consécutifs pour éviter une trop grande corrélation des données.

La forme d'enregistrement la plus simple consiste à reporter les types et horaires de consommation d'aliments et boissons sans détail sur les quantités. Ce type d'enregistrement peut être utile pour déterminer des profils de consommation, mais ne permet pas d'estimer de manière précise les apports.

Plus fréquemment, il est également demandé au sujet de préciser les quantités consommées. Les estimations les plus précises sont obtenues par la pesée directe des aliments à l'aide d'une balance. Cette technique nécessite de la part du répondant une coopération et un investissement très importants qui risquent de renforcer l'apparition de certains biais liés à la méthode d'enregistrement.

Peser les aliments donne au sujet l'opportunité de prendre concrètement conscience de ses apports et risque ainsi d'influencer ses consommations pendant la période

d'enregistrement. Les données recueillies ne reflèteront donc pas les consommations habituelles. La méthode d'enregistrement nécessite de savoir lire et écrire et sa lourdeur risque de sélectionner la population la plus motivée.

La quantification en unités ménagères (cuillère, bol, verre...) préalablement calibrées par l'enquêteur, ou la présentation au répondant de modèles de photographies sont d'autres moyens couramment employés pour estimer les quantités consommées. L'influence sur les habitudes alimentaires est potentiellement moins importante qu'avec la méthode par pesée.

Une autre alternative est de demander au répondant de photographier les aliments ou les repas avant de les consommer, la quantification étant alors laissée à l'appréciation de l'enquêteur.

Afin d'obtenir une bonne qualité de données à partir de ce type de recueil, il est nécessaire de former les participants pour la description précise des aliments (noms, préparations, ajout de condiments, prise en compte des snacks, etc.) et l'estimation des quantités. Théoriquement, l'enregistrement est fait en temps réel au moment de la prise alimentaire, mais des dictaphones peuvent être utilisés pour faciliter le recueil, en particulier chez les sujets peu lettrés. Chez les enfants, l'enregistrement peut éventuellement être réalisé par une tierce personne. À la fin de l'enregistrement, un enquêteur entraîné revoit avec le répondant l'ensemble des données afin de les clarifier et de rechercher d'éventuels oublis.

Rappel des 24 heures

Le rappel des 24 heures est réalisé au cours d'un entretien pendant lequel on demande au sujet de se remémorer et de décrire tous les aliments et boissons consommés pendant les 24 h précédentes. L'entretien peut se faire en face-à-face ou par téléphone, avec des résultats comparables.

Par son interrogatoire, l'enquêteur a pour rôle d'aider le répondant à rapporter ses consommations, tout en évitant de l'influencer dans ses réponses. Sa formation et sa compétence sont donc primordiales. Le rappel, généralement fait selon l'ordre chronologique des prises alimentaires de la veille, est affecté par les défauts de mémorisation du répondant.

Une technique a été développée aux États-Unis pour améliorer la qualité du rappel et limiter la sous-déclaration des répondants. Dans cette technique, l'interrogatoire est

guidé par une série de questions qui portent spécifiquement sur certains points source d'erreurs ou d'oublis. Ce rappel est dit « à passages multiples », parce qu'il est réalisé en 5 étapes successives :

- 1) « la liste rapide », étape dans laquelle il est demandé au répondant de se souvenir des aliments et boissons consommés la veille de l'entretien, en utilisant sa propre méthode de rappel ;
- 2) « la liste des oublis » au cours de laquelle l'enquêteur interroge le répondant sur les consommations connues pour être fréquemment oubliées (sucreries, snacks, boissons alcoolisées ou non...);
- 3) « les horaires et occasions » des différentes consommations sont ensuite renseignés ;
- 4) « le passage détaillé » a pour but de faire préciser au répondant, à l'aide de questions et d'outils standardisés, chacune de ses consommations et d'en évaluer les quantités. Les lieux de consommation et la durée séparant les prises alimentaires sont également indiqués ;
- 5) la dernière étape consiste à passer en revue l'ensemble des réponses qui peuvent être complétées si besoin.

7.3.2. Recueil des apports habituels

Histoire alimentaire

Contrairement aux méthodes précédentes qui évaluent les apports alimentaires sur une période précise, l'histoire alimentaire cherche à évaluer les habitudes alimentaires typiques du sujet.

La méthode, initialement décrite dans les années 40, comprenait plusieurs étapes dont un enregistrement alimentaire de 3 jours destiné à vérifier les données recueillies lors de l'entretien. En pratique, l'enregistrement alimentaire n'est que très rarement réalisé.

Pendant l'entretien, l'enquêteur interroge dans le détail le répondant sur la répartition habituelle de son alimentation afin d'apprécier son profil alimentaire. Cependant, les apports alimentaires variant dans le temps, il est difficile de définir un profil alimentaire typique sans définir une période de temps à laquelle il se rapporte. Ainsi, en fonction des objectifs de l'enquête et de la typologie du répondant, l'interrogatoire pourra porter sur une période variable correspondant, par exemple, à

une semaine typique, une quinzaine typique, une saison typique voire à une période précise de la vie.

Pour faciliter le rappel, l'histoire alimentaire est souvent retracée en fonction des repas. Mais l'approche basée sur les repas n'est pas la mieux appropriée chez les sujets, de plus en plus nombreux, pour qui les prises alimentaires ne sont plus rythmées par les repas classiques. Elle risque dans ce cas d'entraîner une sous-estimation des apports et une mauvaise évaluation du profil alimentaire en écartant du recueil les consommations interprandiales.

Un rappel des 24 heures bien conduit peut être utile pour débiter l'entretien. Les consommations de la veille pourront ainsi servir de base à l'étude des variations habituelles de consommations (catégories d'aliments, composition des repas et répartition des prises alimentaires).

Concernant l'évaluation des quantités, une autoappréciation qualitative est souvent suffisante (apports élevés, moyens, faibles), mais des informations plus précises peuvent être obtenues à l'aide d'outils spécifiques (modèles, photographies, unités ménagères).

Une histoire alimentaire demande en général au moins 1 heure d'entretien, et nécessite, comme pour le rappel des 24 h, un enquêteur particulièrement entraîné à orienter le répondant par des questions précises, mais toujours neutres.

Il n'en reste pas moins que, comme pour le rappel des 24 h mais de manière plus marquée avec l'histoire alimentaire, les données obtenues avec ce type d'enquête sont très liées au répondant et aux compétences de l'enquêteur. La comparaison des résultats entre individus peut être plus délicate qu'avec d'autres méthodes.

Questionnaires de fréquence de consommation

Les questionnaires de fréquence sont utilisés pour évaluer la consommation habituelle de certains aliments. Il s'agit de la méthode d'enquête alimentaire la plus simple d'utilisation, mais aussi probablement celle qui demande le plus gros travail de préparation en amont.

Un questionnaire de fréquence est constitué d'une liste d'aliments auxquels sont associées des catégories de fréquence de consommation (en nombre de fois par jours, par semaine, par mois, etc.). Il est demandé au répondant de cocher, pour chaque aliment de la liste, la fréquence qui s'approche le plus de sa consommation habituelle.

Le choix ou la création d'un questionnaire de fréquence dépend avant tout de la population ciblée et de l'objectif de l'enquête qui peut être d'évaluer la consommation d'aliments, de catégories d'aliments ou de nutriments.

Lorsque le temps et les moyens sont limités, il est possible d'adapter un questionnaire déjà existant, après s'être assuré de la pertinence de son utilisation. Pour cela, il est important de savoir si le questionnaire original a été validé, pour quel objectif, pour quelle population et à quelle époque il a été créé.

Lorsque le temps et les moyens le permettent, la création d'un questionnaire spécifique aux besoins de l'enquête est préférable, mais demande une méthodologie rigoureuse.

La variabilité des choix alimentaires d'un groupe d'individus (aliments, marques, modes de préparation, etc.) est très vaste et ne peut pas être représentée de manière exhaustive dans un questionnaire de fréquence. Le choix des items alimentaires à inclure dans la liste est ainsi crucial pour le succès du questionnaire.

En règle générale, pour qu'un item alimentaire soit informatif au sein d'un questionnaire de fréquence, il doit répondre à 3 critères :

- ✓ Être consommé assez fréquemment par un nombre non négligeable de sujets.
- ✓ Contenir en quantité suffisante le nutriment/l'aliment dont l'apport est étudié.
- ✓ Mais aussi être consommé en quantité (fréquence) variable selon les individus pour que le questionnaire soit discriminant.

La connaissance préalable des habitudes alimentaires de la population étudiée, soit par la réalisation d'une autre enquête alimentaire, soit par l'exploitation de données existantes récentes est donc indispensable pour choisir avec pertinence les items à inclure dans la liste. La sélection se fait à partir de ces données grâce à des méthodes statistiques de régression pour retenir les aliments qui contribuent le plus aux apports du nutriment étudié et qui permettent de classer les individus en fonction de leur consommation.

Le nombre d'items à retenir est un autre point méthodologique à ne pas négliger. La liste d'aliments peut varier de quelques items à quelques centaines d'items. La tentation serait de construire le questionnaire le plus précis possible pour obtenir un grand nombre d'informations. Or, la coopération du répondant et la précision de ses réponses diminuent avec la longueur du questionnaire. De plus, il a été montré que le gain de précision obtenu par l'accroissement du nombre d'items décroît rapidement avec

l'allongement du questionnaire. La longueur de la liste est en fait déterminée par l'objectif du questionnaire. En général, plus l'objectif est spécifique, plus le questionnaire a tendance à être concis. Par exemple, un questionnaire évaluant l'apport en folates ou en phytoestrogènes sera plus court qu'un questionnaire évaluant les apports énergétiques totaux.

Un questionnaire dont le but est d'évaluer de manière absolue le niveau d'apport d'un nutriment donné sera plus long qu'un questionnaire utilisé pour simplement dépister les « grands » ou « petits » mangeurs au sein d'une population. Lorsque l'on cherche à estimer quantitativement les apports, il faut savoir que la longueur de la liste peut influencer les résultats : les listes longues ont tendance à surestimer alors que les listes courtes ont tendance à sous-estimer les apports.

Une fois la liste établie, la dernière étape de la création du questionnaire est d'obtenir une mesure de la fréquence de consommation qui peut être complétée d'une information sur la taille des portions (questionnaires semi-quantitatifs). Là encore, le nombre de propositions dépend de l'objectif de l'enquête et de la population étudiée, mais les catégories de fréquence devraient toujours être continues, sans « trous », afin que chaque répondant puisse trouver la catégorie qui correspond le mieux à sa consommation habituelle.

7.3.3. Recueil de données au niveau collectif

À l'échelle d'un pays : données de disponibilité alimentaire

Les bilans de disponibilité alimentaire d'un pays fournissent la quantité d'aliments disponible sur le marché intérieur à une période donnée, exprimée en poids ou volume par habitant et par jour. Ces données sont basées sur les statistiques agricoles nationales et internationales (FAO, OCDE, Eurostat). Elles sont obtenues par la somme de la production nationale et des importations à laquelle sont retranchées les exportations, les pertes (ou déchets) et les utilisations alimentaires animales.

Ces statistiques nationales, calculées sur les mêmes bases année après année, sont intéressantes pour suivre l'évolution des grandes tendances de consommations d'un pays. Cependant, elles restent assez grossières puisque fondées sur les seuls flux économiques. Ramenées au nombre d'habitants, elles sont largement surestimées car les pertes aux différents stades, depuis la production jusqu'à l'assiette du consommateur, ne sont pas évaluées. Par ailleurs, il s'agit de moyennes nationales qui ne tiennent pas

compte de différents facteurs liés à l'alimentation comme l'âge, le sexe ou le niveau socioculturel. Ainsi, ces données doivent être prises pour ce qu'elles sont, soit des données de disponibilités alimentaires totales. Elles ne peuvent pas être assimilées aux consommations totales, et encore moins aux consommations individuelles.

7.4. Exploitation des données

Les données de consommations recueillies au niveau de l'individu peuvent être exploitées selon 2 types d'approches. Par ailleurs, en fonction du contexte de réalisation de l'enquête (épidémiologique ou clinique), les résultats seront également exploités différemment.

Deux types d'approches : aliment/nutriment vs profil alimentaire

L'exploitation des données peut être abordée sous 2 approches :

- ✓ soit par l'étude de la consommation isolée d'un aliment particulier ou d'un nutriment (après conversion des données par l'intermédiaire d'une table de composition des aliments) ;
- ✓ soit par l'étude de l'alimentation dans sa globalité (ou « profil alimentaire »).

En épidémiologie nutritionnelle, le premier type d'approche a permis de comprendre de nombreuses relations physiopathologiques unissant l'alimentation et certains processus pathogènes. Toutefois, l'étude isolée d'un nutriment présente des limites liées à la complexité de l'exposition alimentaire. Les aliments contiennent par essence de nombreux nutriments, il est donc difficile d'analyser leurs effets propres. De plus, les nutriments interagissent entre eux dans le bol alimentaire, ce qui peut influencer leur biodisponibilité et leur absorption. Par conséquent, la mesure de l'apport alimentaire d'un nutriment donné a peu de chance de correspondre à la quantité de nutriment qui sera disponible métaboliquement.

Enfin, les enquêtes nutritionnelles montrent que les consommations de certaines catégories d'aliments sont corrélées entre elles pour diverses raisons culturelles, saisonnières ou de contrainte de production locale. Par exemple, les consommateurs de fruits et légumes sont généralement des consommateurs de poissons. Un consommateur d'huile d'olive est souvent un mangeur de légumes. Dans ces conditions, il apparaît difficile de discerner la contribution de l'un ou l'autre aliment (nutriment) à l'état de santé des individus.

Dans ce contexte, l'étude de l'alimentation dans sa globalité, ou des « profils alimentaires », apporte un point de vue différent en considérant que la consommation d'un aliment ou d'un nutriment ne peut être dissociée du reste de l'alimentation d'un individu.

En épidémiologie, des méthodes de scoring (score de variété, score de diversité, score de qualité globale de l'alimentation) ou des techniques de modélisations statistiques (analyses en composante principale, analyses en cluster) se sont largement développées pour appréhender l'étude des profils alimentaires.

Deux contextes : épidémiologie vs clinique

L'exploitation des données sera bien évidemment différente selon qu'il s'agit d'un recueil de l'apport alimentaire réalisé dans le cadre d'un entretien clinique ou dans le cadre d'une enquête épidémiologique.

En épidémiologie, 4 grands types d'analyses peuvent être individualisés, chacun répondant à des questions spécifiques :

- 1) Estimation des apports moyens d'une population pour les comparer à d'autres populations ou en fonction de certaines caractéristiques de la population ;
- 2) Estimation de la distribution des apports dans la population d'intérêt pour étudier la proportion de sujets ayant des apports excessifs ou insuffisants ;
- 3) Analyses de corrélations ou de régressions pour rechercher d'éventuelles relations entre le niveau d'apport alimentaire et d'autres mesures (niveau de pression artérielle, indice de masse corporelle, cholestérolémie, par exemple) ;
- 4) Analyses qualitatives dans lesquelles les sujets sont classés en catégories selon leurs apports alimentaires (les catégories sont définies par des intervalles fixes) et la variable étudiée (le plus souvent en rapport avec l'état de santé : présence ou non d'un cancer, d'une obésité, d'un syndrome métabolique, par exemple).

On remarque qu'en épidémiologie, en fonction de la question posée, la précision et l'exactitude des données auront une importance et des conséquences variables. L'exactitude des données est finalement peu importante dans les analyses de corrélations, par exemple, où l'essentiel est d'obtenir un classement correct des sujets selon leurs apports. Au contraire, dans les études de distribution ou pour l'estimation des apports moyens, la qualité des résultats et des conclusions est très dépendante de l'exactitude

des données. Ici se trouve tout l'enjeu des méthodes de correction statistiques que nous évoquerons par la suite.

En clinique, le problème est différent. Les méthodes de recueil des consommations décrites ci-dessus, souvent longues et contraignantes, pourront être adaptées ou appliquées plus simplement. Dans ce contexte, les enquêtes alimentaires sont avant tout des outils d'aide à l'analyse du comportement alimentaire global du patient. Elles permettent d'initier un dialogue autour de l'alimentation et de repérer d'éventuels troubles du comportement alimentaire. L'analyse du profil alimentaire est la première étape de la prescription diététique et est essentielle pour le suivi du patient. Les informations les plus importantes à recueillir portent sur les goûts et préférences alimentaires, les catégories d'aliments les plus fréquemment consommés et les circonstances dans lesquelles ils sont consommés. Le décompte exact des calories ne sera finalement que très rarement nécessaire.

7.5. Limites des enquêtes alimentaires

Quelle que soit la méthode utilisée, l'enquête alimentaire est toujours soumise à des erreurs. Il ne faut pas pour autant renoncer aux enquêtes alimentaires ni en conclure qu'il s'agit de mauvais instruments, mais plutôt comprendre que l'identification des erreurs est importante pour l'interprétation des résultats.

La qualité d'une enquête alimentaire est déterminée par sa précision et sa validité.

Précision (ou reproductibilité) de la méthode

Une méthode précise est une méthode reproductible, c'est-à-dire une méthode capable de donner des résultats comparables lorsqu'elle est reproduite sur un même échantillon et dans les mêmes conditions expérimentales. Autrement dit, la précision d'une méthode est une estimation de la dispersion des valeurs obtenues dans les mêmes conditions d'étude. Il est important de ne pas confondre la notion de précision (ou reproductibilité) avec la notion d'exactitude. En effet, une méthode peut être très précise (ou reproductible) sans pour autant estimer de manière exacte les apports alimentaires. La mesure du coefficient de variation entre 2 répétitions de la méthode permet d'estimer le degré de précision (ou reproductibilité).

Validité (ou exactitude) de la méthode

Une enquête alimentaire valide est une enquête qui estime de manière exacte les apports alimentaires réels sur la période d'observation déterminée. Par exemple, un enregistrement alimentaire est valide s'il permet de recueillir de manière exacte toutes les consommations (aliments ou boissons) du sujet ayant eu lieu pendant la période d'enregistrement. Là encore il convient de distinguer la notion de validité d'un enregistrement de la notion de représentativité des apports habituels. En effet, un enregistrement peut être tout à fait valide sans pour autant correspondre aux consommations habituelles du sujet, en particulier lorsque la méthode utilisée pour recueillir les apports a eu pour effet de modifier, consciemment ou non, les habitudes alimentaires du sujet enquêté.

Pour être valide et représentative de l'alimentation habituelle, une enquête alimentaire devrait refléter ce que le sujet aurait consommé s'il n'avait pas été enquêté. Nous verrons plus loin qu'il existe des méthodes de validation pour vérifier l'exactitude de la méthode.

Une méthode idéale serait une méthode permettant de renseigner sur les consommations des individus avec une précision et une exactitude irréprochables. Cela supposerait qu'il n'y ait aucune erreur dans le recueil, l'analyse et l'interprétation des données. Assurément, aucune méthode d'enquête alimentaire ne remplit les conditions d'une méthode idéale. Quelle que soit la méthode employée, sa précision et son exactitude seront affectées par un certain degré d'erreur.

Nature de l'erreur

Les erreurs sont classiquement classées en 2 catégories : les erreurs randomisées (ou aléatoires), et les erreurs systématiques.

Les erreurs randomisées sont des erreurs attribuables au hasard. La présence d'erreurs aléatoires augmente la variabilité des estimations et par conséquent diminuent la précision de l'enquête. La précision des estimations étant déterminée par la taille de l'échantillon et la variabilité du paramètre étudié, ce genre d'erreur peut être limité dans une enquête alimentaire en augmentant soit le nombre d'individus observés, soit la durée d'observation de chaque individu.

Les erreurs systématiques, par contre, ne sont pas randomisées et ne peuvent donc pas être atténuées en augmentant le nombre d'observations. Il faut alors chercher à les

éviter au maximum car elles sont difficiles à corriger a posteriori. Il s'agit, par exemple, d'erreurs liées à la méthode (questionnaire ou table de composition inappropriée). Ce genre d'erreur conduit à l'apparition de biais dans l'estimation des apports alimentaires et altère la validité de l'enquête.

En épidémiologie, on distingue 3 types de biais qui se produisent pendant :

- 1) La sélection des sujets = biais de sélection (les sujets inclus dans l'étude ne constituent pas un groupe représentatif de la population ciblée).
- 2) L'observation des sujets = biais d'observation ou de classement (résultant d'erreurs commises systématiquement dans la collecte des informations).
- 3) L'analyse des données = biais de confusion (lorsque la relation mise en évidence par l'analyse est en partie liée à l'existence d'un tiers facteur).

L'impact de chaque type d'erreur sera différent selon la nature de la question posée et la méthode d'analyse utilisée pour y répondre.

Les différentes sources d'erreurs dans les enquêtes alimentaires

La nature et le degré de l'erreur dépendent à la fois de la méthodologie utilisée et des sujets étudiés. Les sources d'erreurs sont très nombreuses au cours d'une enquête alimentaire, et la distinction théorique entre erreur randomisée et erreur systématique n'est pas toujours évidente. La Figure ci-dessous résume les différentes sources d'erreurs rencontrées au cours d'une enquête alimentaire.

→ Variabilité des apports alimentaires (erreur randomisée)

Les apports alimentaires des individus ne sont pas stables dans le temps. Ils varient de manière qualitative et quantitative d'un jour à l'autre, d'une semaine à l'autre voire d'une année à l'autre. En général, le recueil des apports alimentaires ne peut se faire que sur de courtes périodes et ne peut pas refléter les apports habituels (au long cours) des individus.

Lorsque l'objectif est d'estimer les apports habituels des sujets, pour mettre en évidence une relation avec une pathologie chronique, par exemple, la variabilité journalière des apports peut être considérée comme une erreur au sens statistique du terme (par opposition aux erreurs de mesure).

→Erreurs liées à la table de composition (erreur systématique)

Pour convertir avec la meilleure précision possible les données des aliments aux nutriments, il faut disposer d'une table de composition de bonne qualité et adaptée à la méthode d'enquête.

Pour être de bonne qualité, une table de composition doit répondre à plusieurs critères:

- 1) Être récente ou mise à jour régulièrement en raison de l'apparition constante de nouveaux aliments, des changements fréquents de composition des aliments, et de l'évolution des techniques d'analyse de leur composition ;
- 2) Être adaptée à la population étudiée (les aliments analysés doivent correspondre aux aliments consommés par la population étudiée) et au type d'enquête ;
- 3) Être précise dans la description des aliments car la composition en nutriments est variable selon le mode de consommation (cru ou cuit, avec ou sans déchet, etc.) ;
- 4) Être précise dans l'estimation de la composition en nutriments de chaque aliment ;
- 5) Être la plus complète possible pour limiter au maximum le nombre de données manquantes sur les nutriments de manière à ne pas sous-estimer les apports.

→Erreurs d'estimation des quantités

La mesure quantitative des apports alimentaires est déterminée par la mesure correcte de la fréquence de consommation et des quantités consommées. Différents outils permettent, en dehors de la pesée directe des aliments, d'aider l'individu à estimer les quantités consommées: des outils en 3 dimensions comme les modèles d'aliments, ou en 2 dimensions comme les photographies, les dessins, les simulations graphiques informatiques, etc. Cependant, l'estimation des tailles de portions représente une source majeure d'erreur dans le recueil de données.

L'identification des tailles de portion est un processus complexe dans lequel les facultés de perception, de conceptualisation et de mémorisation jouent un rôle important.

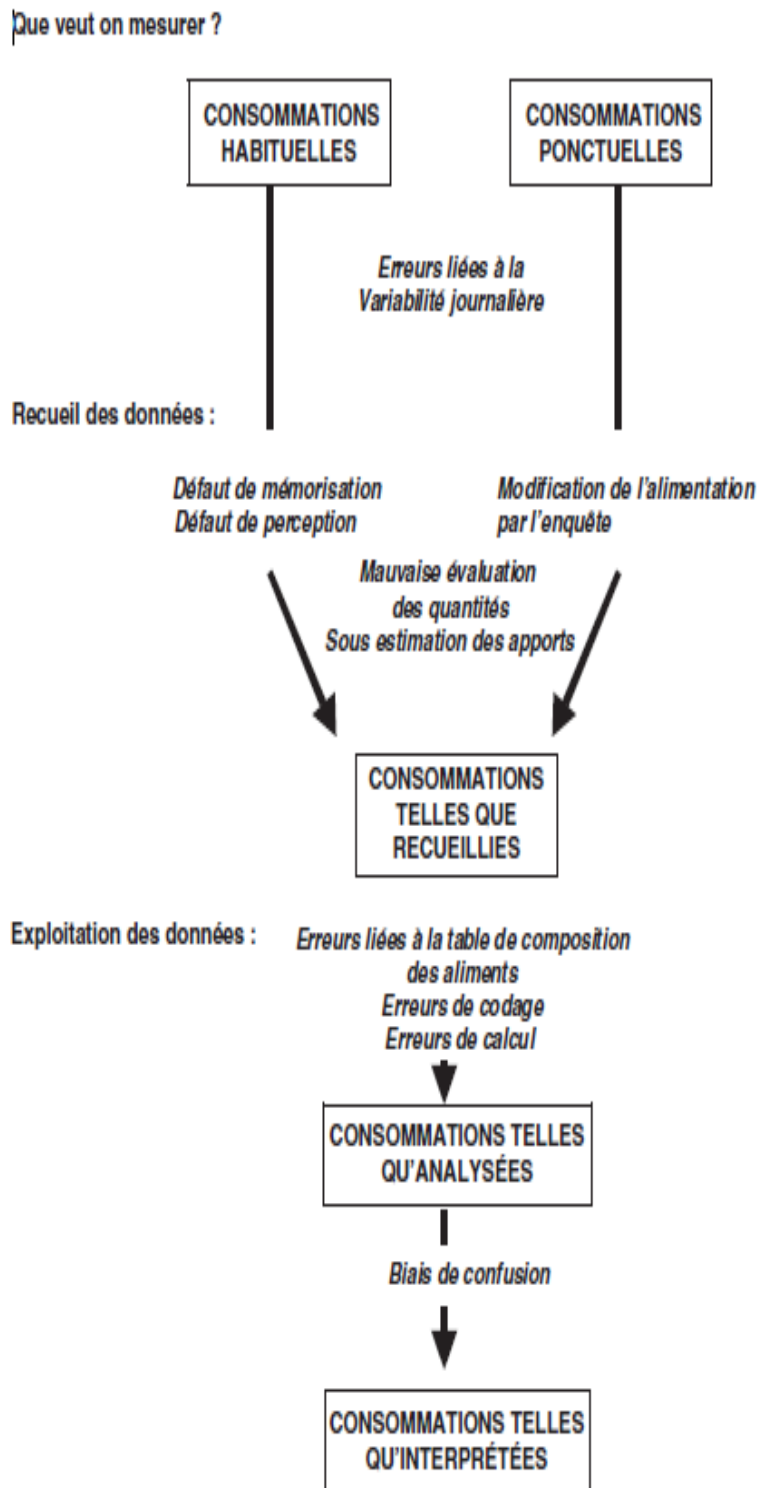


Figure : Principales sources d'erreurs rencontrées aux différentes étapes de l'enquête alimentaire.

RÉFÉRENCES BIBLIOGRAPHIQUES



RÉFÉRENCES BIBLIOGRAPHIQUES

Dagnélie, P. (2004). Principes d'expérimentation-diffusion, Lavoisier, Paris.

Grais, B. (2003). Méthodes statistiques, Dunod.

Houdret, J.L. et Malherbe, L.(2005). Méthodes de reconstitution de moyennes et de nombres de dépassements de seuils a partir de données de campagnes. Rapport LCSQA disponible à l'adresse www.lcsqa.org.

Kherri, A. (2011). Support pédagogique de cours N° 02 : Echantillonnage. www.sg-ehedc.jimdo.com. 20p.

Lebart, L., Morineau, A. et Fenelon J.P. (1982). Traitement des données statistiques, seconde édition, Dunod.

Lecoutre, J. P. (2006). Statistique et probabilités : manuel et exercices corrigés, Dunod.

Norme ISO 9359:1989. Qualité de l'air - Echantillonnage stratifié e pour l'estimation de la qualité de l'air ambiant.

Saporta, G. (2006). Probabilités, analyse des données et statistique 2e édition révisée et augmentée, Technip.

Tille, Y. (2001). Théorie des sondages. Echantillonnage et estimation en populations finies. Cours et exercices avec solutions, Dunod.

Tomassone, R., Audrain, S., Lesquoy-Deturckheim, E. et Millier, C. (1992). La régression : Nouveau regard sur une ancienne méthode statistique seconde édition, Masson.

Vesseau, A. (1988). Méthodes statistiques en biologie et en agronomie. Tech et Doc Lavoisier, Paris.

Viliam, M. (1999). Méthodes expérimentales en agronomie. Tech et Doc Lavoisier, Paris.

Site web:

<http://www.iutbayonne.univ-pau.fr/~grau/2A/stat/cadre2.html>

