



Titre : Quelle population étudier ?

Intervenant : Marthe-Aline Jutand

Bonjour à vous tous,

Je suis ravie de vous retrouver aujourd'hui pour aborder un sujet qui me tient tout particulièrement à cœur en tant que statisticienne : il s'agit de la définition de la population que l'on souhaite enquêter et de la construction de l'échantillon.

Nous débutons cette séquence par quelques définitions.

Nous définirons une **population** comme un ensemble d'éléments (êtres humains, êtres vivants ou objets) soumis à une étude statistique. Si l'on souhaite réaliser une enquête sur la qualité des dossiers patients nous parlerons, par exemple, de la population de tous les dossiers patients du Centre Hospitalier Universitaire de Bordeaux pour l'année 2016 ; si on souhaite réaliser une enquête sur les soins bucco-dentaires chez les enfants de moins de 11 ans, nous nous intéresserons, par exemple, à la population de tous les enfants de moins de 11 ans scolarisés en Gironde.

Une **unité statistique** quant à elle correspond à tout élément faisant partie de la population visée, pour lequel des données sont recueillies.

Pour introduire deux définitions importantes que sont la **population-cible** et la **population-source**, nous allons reprendre la définition d'un objectif et insister sur l'importance de bien préciser dans ce cadre le temps, le lieu mais aussi les personnes. Ces trois éléments définissent la population que l'on souhaite étudier dans le cadre de l'enquête. C'est ainsi que nous déterminerons la population cible.

Nous reprenons l'exemple de l'étude nationale nutrition santé (ENNS), dont l'objectif principal est de décrire les apports alimentaires et l'état nutritionnel des adultes et des enfants vivant en France métropolitaine en 2006-2007. La population-cible est composée des adultes et des enfants vivant en France métropolitaine en 2006-2007.

La population-cible correspond donc à la population qui a motivé au départ la mise en place de l'enquête et que nous voudrions atteindre pour obtenir l'information nous permettant de répondre à notre objectif.

Mais malheureusement, il y a souvent un écart entre ce que l'on veut et ce que l'on peut.

Dans le cadre de l'étude nationale nutrition santé, le recrutement des sujets ainsi que les recueils de données ont été réalisés en différentes phases et il est important d'en prendre connaissance pour définir la population atteignable. Nous empruntons une figure de l'article « Etude nationale nutrition santé : situation nutritionnelle en France en 2006 selon les indicateurs d'objectif et les repères du Programme national nutrition santé », publié en 2007 par l'USEN (unité de surveillance et d'épidémiologie nutritionnelle) présentant le déroulement de l'étude. Il y a eu tout d'abord un envoi de

brochure, puis un contact téléphonique et une visite à domicile. La totalité de la population-cible initialement recherchée ne pouvait pas être atteinte compte tenu de ces contraintes. Les unités pouvant être incluses dans l'enquête devaient être accessibles par téléphone et avoir un domicile. Ce qui excluait de fait les personnes sans domicile fixe par exemple. La population joignable était donc les adultes et enfants vivant en France métropolitaine en 2006-2007 ayant un domicile fixe et joignables par téléphone.

Cette population est appelée population-source et est un sous-ensemble de la population-cible. Nous définirons donc la population-source comme étant la population accessible lors de l'enquête selon la stratégie définie. C'est donc au sein de cette population-source que les participants de l'enquête (c'est-à-dire l'échantillon) seront sélectionnés. Nous vous proposons un moyen mémo-technique pour vous rappeler cela : « l'échantillon est puisé dans la population-source ».

On peut étudier l'écart entre la population-cible et la population-source en décrivant **la couverture de la population-cible par la population-source**. Cette notion de couverture correspond à deux critères, un **critère quantitatif** et un **critère qualitatif** :

- le critère quantitatif correspond à la proportion que représente l'effectif de la population-source par rapport à l'effectif de la population-cible
- le critère qualitatif correspond au caractère de ressemblance ou de dissemblance des unités appartenant à la population-cible et non à la population-source par rapport aux unités appartenant aux deux populations.

Si l'on considère une enquête dont le sujet est d'« estimer la prévalence des caries dentaires chez les enfants des écoles primaires de Gironde en décembre 2016 », la population-cible correspond donc à tous les enfants inscrits dans une école primaire de Gironde en décembre 2016. S'il a été prévu de réaliser une enquête avec un examen dentaire réalisé dans les écoles, la population-source se réduit donc à tous les enfants présents dans les écoles primaires en décembre 2016 et dont les parents auront donné leur autorisation pour que leur enfant ait cet examen dentaire. La question de la couverture peut se décliner de la manière suivante:

- *Le critère quantitatif devrait être assez bon, c'est-à-dire correspondre à un pourcentage élevé, car la proportion d'enfants malades, ou hospitalisés ou absents pour tout autre raison, ne devrait pas être importante (sauf événement exceptionnel), cependant il faudra aussi soustraire la proportion d'enfants dont les parents ne donneraient pas l'autorisation. Pour réduire cette proportion, les responsables de l'enquête devront mettre tout en œuvre pour une bonne communication permettant de sensibiliser les parents sur l'importance de l'enquête.*
- *Quant au critère qualitatif, il est difficile de savoir si les indicateurs qui intéressent les responsables de cette enquête diffèrent entre la population-source et l'ensemble des enfants qui ne seront pas inclus. On peut supposer que l'absence d'un enfant de l'école durant le mois de décembre ne sera pas due à son état bucco-dentaire ; par contre, il est plus difficile de savoir si la non-autorisation parentale ne serait pas due à des mauvais soins bucco-dentaires.*

Pour obtenir une bonne couverture de la population-cible par la population-source, il sera nécessaire de réaliser une bonne communication d'informations pour ne pas inquiéter les parents et les inciter à donner leurs autorisations de participation.

Intéressons-nous maintenant à la question de la **sélection**. Dès lors que la population-source est définie, deux issues sont possibles pour obtenir les unités qui seront incluses dans l'enquête :

- soit la totalité des unités statistiques de la population-source peuvent être incluses dans l'enquête et il s'agit d'un recensement,
- soit un sous-ensemble de la population source sera sélectionné pour participer à l'enquête et il s'agira d'un **échantillonnage**

Quels sont les critères qui permettent de choisir entre la mise en œuvre d'un échantillonnage ou la réalisation d'un recensement ?

- L'élément le plus important pour ce choix est la taille de la population source. Plus cette taille est importante, plus il est difficile de faire un recensement car le coût de réalisation de l'enquête sur l'ensemble de la population serait très élevé pour avoir des informations de bonne qualité.
- Un autre élément important est la durée de l'enquête. Plus l'enquête doit être réalisée dans une période courte, moins il est réalisable de faire un recensement.

On définit l'échantillon comme étant un sous-ensemble d'unités statistiques de la population-source, et l'échantillonnage comme étant le processus de sélection d'un échantillon au sein de la population-source.

En résumé, nous considérons la population-cible comme étant la population définie initialement par l'objectif, appelée parfois population contextuelle : c'est la population que l'on veut étudier. La population-source est quant à elle la population que l'on peut atteindre au regard de l'organisation de l'enquête : c'est un sous-ensemble de la population-cible qui couvre plus ou moins bien la population-cible. Enfin, on s'interroge sur la possibilité ou non de réaliser un recensement de la population-source ; si cela n'est pas possible, on réalisera un échantillon, en sélectionnant un sous-ensemble d'unités de la population-source.

Nous terminons donc ainsi la présentation du diaporama « Quelle population étudier ? ».