

TECHNIQUES D'ENQUÊTE QUALITATIVE ET QUANTITATIVE

Rifat Haxhijaj – rifat.haxhijaj@unifr.ch

Séance 3 – 06 mars 2017

L'échantillonnage

"Qui interroger" suit la construction de la problématique, de la question de recherche, des hypothèses et est la dernière étape avant d'aller sur le terrain.

« L'échantillon ne doit être précisé que lorsque l'objet de l'enquête est suffisamment défini, construit. En effet, une partie des hypothèses peut contribuer à définir les personnes à interroger » (François de Singly, 2005, p. 41)

- Quelle est la population qu'il est nécessaire de connaître?
- Comment choisir dans cette population les personnes à interroger?

Recensement: On fait un recensement pour dénombrer tous les individus ou éléments d'une population et pour en connaître un certain nombre de caractéristiques.

- Mais faire des recensements n'est pas toujours pratique;
- Coûteux et trop long
- Avoir un bon échantillon permet de "représenter" l'ensemble de la population qui nous intéresse.

Rappel de notions

- Population: ensemble des éléments sur lesquels porte l'enquête (peut être des institutions, des personnes, des articles...)
- Échantillon: partie de la population sélectionner pour mener l'enquête
- Échantillonnage: action de recueillir de l'information sur une fraction (échantillon) de l'ensemble (population) que nous voulons étudier.

Rappel; logiques du quantitatif

4 logiques à la base de toute recherche quantitative:

1. Logique de quantification de la réalité sociale
2. Logique de généralisation des données
3. Logique de réplication de la recherche
4. Logique de causalité

Types de données

- Données secondaires (stats faites par d'autres)
 - Données issues de fichiers administratifs
 - Données issues des recensement
- Données primaires (stats provenant de votre enquête)
 - Données recueillies au moyen de votre questionnaire
 - Données textuelles et d'observation directe (qui repose sur le comptage)

Choisir un échantillon représentatif !

Il est inutile d'interviewer l'ensemble de la population puisqu'un échantillon, si correctement choisi, représente l'ensemble de la population.

Échantillon représentatif

- Un échantillon est dit représentatif quand il possède les mêmes caractéristiques que la population que l'on souhaite étudier.
- Faute de représentativité, les résultats obtenus ne peuvent pas être généralisés à la population.
- On ne cherche pas toujours la représentativité. L'accès aux données et la commodité peuvent surpasser la volonté de représentativité.

Deux grands types d'échantillons

- Échantillon représentatif (probabiliste): utilisé lors des enquêtes scientifiques rigoureuses
- Échantillon non-représentatif (non-probabiliste): soit lorsque les chercheurs testent leur instrument, soit dans le cadre des enquêtes non-scientifiques (médias, journalistes, vox-pop, combat des clips, etc.)

...deux grands types d'échantillons

- **L'échantillon probabiliste (représentatif, aléatoire):**
 - ✓ chaque personne dans la population a une chance (probabilité) égale, ou au moins connue, d'être choisi pour faire partie de l'échantillon
- **L'échantillon non-probabiliste (non-représentatif, raisonné).**
 - ✓ certaines personnes, certains individus ont une chance plus grande, mais inconnue que d'autres d'être sélectionnés.

...deux grands types d'échantillons

L'échantillon probabiliste (représentatif, aléatoire)	L'échantillon non-probabiliste (non-représentatif, raisonné)
<ul style="list-style-type: none">• Aléatoire simple• Systématique• Stratifié• Par grappes	<ul style="list-style-type: none">• Par quota• Boule de neige• Commodité (aveuglette)• Volontaire

A. Technique probabiliste

- Manière la plus sûre d'offrir à tous la même chance de sélection, par choix aléatoire, puisque tous les éléments de la population ont la même chance d'être choisis.
- Implique d'énumérer tous les membres de la population puis de tirer leur nom au hasard "hors d'un chapeau"
- Mais impossible de faire cela dans tous les cas...

1a. Échantillon aléatoire simple

- Doit avoir une liste complète de la population qui nous intéresse + une table de nombres aléatoires;
- Il faut numéroter tous les items de la population (ex. si on étudie une pop composé de 5332 étudiants, on les numérote de 0001 à 5332);
- On prend la table de nombres aléatoires, on sélectionne comment on la lit (à partir de la 1ere ligne, 3e colonne), on retient les 4 premiers chiffres, on regarde si ça correspond à un étudiant, si oui on le garde, puis on continue... de gauche à droite;

↓ Point de départ

→ 17031	<u>15532</u>	<u>16006</u>	89840	<u>44231</u>	55053	57859	98136	97770	69560
<u>46068</u>	<u>23960</u>	<u>51491</u>	69217	69235	<u>52827</u>	<u>19263</u>	89194	<u>24726</u>	67945
<u>29287</u>	08655	33171	62907	54727	75276	94979	95023	33087	05835
13452	95111	24993	18476	45482	16698	38247	59071	26588	48372
54822	41406	14108	61813	81646	97975	26103	73433	15984	12248

Les numéros retenus sont soulignés dans l'extrait de la table.

ANNEXE 4

TABLE DES NOMBRES ALÉATOIRES

17031	15532	16006	89840	44231	55053	57859	98136	97770	69560
46068	23960	51491	69217	69235	52827	19263	89194	24726	67945
29287	08655	33171	62907	54727	75276	94979	95023	33087	05835
13452	95111	24993	18476	45482	16698	38247	59071	26588	48372
54822	41406	14108	61813	81646	97975	26103	73433	15984	12248
61157	86506	36521	74559	95909	83124	58410	92789	40796	26484
06297	36767	53523	14324	87207	05735	81259	21006	37041	18175
06380	50714	29468	36676	92934	32241	33596	58725	54268	29531
16922	55536	02285	32303	12394	89836	73410	98591	00932	98470
75522	42808	52924	24688	24178	84663	78611	45748	11730	64919
77676	33246	65833	68819	01858	50522	94887	45814	43572	18939
92087	86889	84083	14650	08095	95590	97963	09335	97474	87457
12320	78625	16733	63782	45620	66638	71798	64974	61135	82106
52356	13378	53650	93061	87111	49382	64647	44638	41368	19047
07901	28539	89636	10385	62725	06591	28382	55094	88209	76728
56343	35970	02447	06844	40971	85887	47863	46252	37497	58570
58144	15741	01028	08871	01297	10732	74618	63693	67541	30104
99547	74293	07796	27114	58187	02955	91212	15016	99462	87670
09414	32790	84727	45600	94054	69968	96711	03620	13178	61944
04277	01667	72142	36426	97145	83220	16067	10268	19617	36639
07364	66305	45505	68845	96583	22892	42200	68326	18363	77033
65558	07846	35533	50124	72186	22898	50040	06093	98906	39580
70782	33400	30616	78710	92467	94177	57873	04405	52316	98513
00265	05446	80843	50570	66276	17289	82648	35778	83999	52254
23675	13661	97400	11151	28082	41771	32730	94989	04214	13908
80817	08512	43918	39234	45514	39559	44774	89216	28606	40950
28426	57962	36149	23071	01274	36321	42625	52913	92319	46602
21699	83058	78365	90291	41255	88742	90959	71911	12364	36500
76196	25126	75243	06455	09699	27259	87118	61898	34070	30663
29910	23090	03597	94026	22920	24888	12060	94269	26857	86418
88434	97038	07584	91689	20272	28419	31588	02629	10304	40897
19631	04682	73840	11159	07635	44424	73856	59622	19881	11576
09207	26807	35382	28345	56942	55028	05253	20412	32181	29668
47677	11154	22639	37745	18850	27436	46074	24186	31239	89847
72402	94986	61941	09619	47521	28566	64866	59213	76003	16418
70120	45782	18412	43100	08253	60693	24399	31778	74211	35853
36947	28111	78904	57756	84812	72046	90147	86847	85179	93120
50512	87468	29926	36505	30657	18856	76880	35595	73896	04326
62112	73064	46212	02202	73464	48659	10954	02926	00794	24981
82353	20212	18866	97787	13678	63246	68120	39120	33821	86064

2a. Échantillon systématique

- Avoir une base de sondage numérotée;
- Choisir la taille de l'échantillon et se fixer un taux:
 - N° de la population/ n° de l'échantillon = intervalle
- Ex. Prof de la fac. lettres $N = 250$, échantillon de 75
 - $250/75 = 3$
- Vous allez donc prendre votre liste et à chaque 3 noms vous allez sélectionner le nom.

3a. Échantillon stratifié

- Lorsque la population est hétérogène, préférable de la découper en strates, de la diviser en groupe (âge, ethnie);
- Chaque item n'appartient qu'à une strate;
- À l'intérieur de chaque strate on échantillonne de manière aléatoire.

- Exemple: on veut savoir les opinions en matière d'environnement des citoyens de Genève, Fribourg et Zurich et on veut s'assurer que des citoyens de chaque canton y figurent;
 - Pour un échantillon de 1000 personnes (parmi une population totale de 595 553 pers.) on devrait interviewer entre autres 311 genevois, et 57 fribourgeois.
 - Si on fait des strates, pour avoir 1000 personnes, on interrogerait 333 genevois, 333 fribourgeois et 333 zurichois.

- Créer des strates permet de mieux comparer des catégories sous-représentées.
- Ex. Femmes cadres, avec échantillon représentatif des femmes actives, les femmes cadres comptent pour 6%. Pour un échantillon de 1000 = 60 femmes cadres, trop peu pour comparer avec les autres professions.
 - Si on établit qu'il y a 6 groupes de professions et qu'on veut un échantillon de 1000 pers. on divise l'échantillon en 6 strates, donc 167 femmes par groupes de professions seraient interviewées.

4a. Échantillon par grappes

- Utilisé lorsque l'on s'intéresse à des groupes d'individus (ménages, entreprises, associations);
- Consiste à tirer au hasard des groupes d'individus, puis soumettre à l'analyse les individus qui composent les grappes;
- Nécessite une population formée par des groupes d'éléments:
 - choisir un échantillon de n sous-groupes
 - questionner tous les différents groupes
 - questionner toutes les personnes faisant partie d'un groupe

- Ex. vous voulez un échantillon puisé parmi tous les étudiants qui diplômeront du secondaire en Suisse. Vous pourriez aller dans toutes les écoles secondaires et demander les listes de noms. Ça sera très long...
 - Choisir aléatoirement une grappe d'écoles secondaires
 - Sélectionner aléatoirement parmi cette grappe d'écoles un échantillon d'étudiants qui diplômeront

Qualité des échantillons probabilistes

- La principale qualité des échantillons probabilistes est que vous pourrez faire des inférences statistiques, c'est-à-dire que vous pourrez généraliser les résultats de votre échantillon à une population générale.

B. Technique non-probabiliste

- C'est l'utilisation d'un échantillon raisonné et non pas aléatoire (donc pas mathématique);
- C'est un échantillon formé par des éléments qui ont une chance plus grande que d'autres d'être sélectionnés.

5b. Échantillon par quota

- C'est la technique qui est la plus utilisée lorsque la technique d'échantillonnage non-probabiliste est employé;
- Elle vise à reproduire un échantillon représentatif sans choix aléatoire;
- Elle se fait par l'utilisation de la proportion;
 - Demande de connaître la structure de la population

- Les intervieweurs doivent trouver des éléments de la pop étudiée en fonction de caractéristiques particulières;
- Demande une sélection de caractéristiques;
- Demande que la répartition de ces caractéristiques soient similaires à l'ensemble de la population;
- Se base donc sur les pourcentages respectifs des différentes catégories (démographique, socio-pro, etc.).

- Si la population qui nous intéresse est composée de 100 000 familles, dont 50 000 familles ont pour personne de référence un ouvrier et 10 000 familles un cadre;
- Si on veut un échantillon de 1 000 familles =
- Interviewer 500 familles d'ouvriers, 100 familles de cadres.

6b. Échantillon boule de neige

- Premier cas de figure: type d'échantillon pour dégager des systèmes de relations existant dans un groupe;
- C'est l'ajout à un noyau d'individus tous ceux qui sont en relation avec eux;
- Permet de comprendre et d'évaluer les comportements, réactions, d'un groupe, dans un contexte donné;
- La méthode consiste à trouver un individu de la population à étudier puis de lui demander de chercher d'autres participants, idem avec les autres participants;

... boule de neige

- Deuxième cas de figure: méthode utilisée pour accéder à des populations peu accessibles ou des groupes dont on a aucun moyen d'entrer en contact avec;
- Idem, vous réussissez à trouver une personne, vous établissez un lien de confiance, vous lui demandez de vous aider à trouver d'autres participants;
 - Puisqu'on suppose qu'elle connaît d'autres personnes qui ont des caractéristiques semblables aux siennes.

7b. Échantillon par commodité (à l'aveuglette, accidentel)

- Ce fait lorsqu'on s'intéresse à une population pour laquelle on a aucune base de sondage, un échantillon pour lequel on recherche des qualités précises, pas représentatives d'un ensemble;
- Et qu'on suppose que les gens qui partagent ces caractéristiques se retrouvent dans certains endroits;
- On aborde alors les gens dans ces endroits, on les sélectionne tous, un sur deux, un sur cinq... On peut donc y aller à l'aveuglette

- On peut être plus rigoureux en créant des échantillons spatial (échantillon aléatoire dans la région étudiée) et temporel (échantillon en fonction des heures e la journée), à l'intérieur desquels on choisira des individus à l'aveuglette;
- On peut être moins rigoureux (comme le sont les médias par exemple) et choisir des personnes à l'aveuglette, sans égard au lieu et au moment de la journée;
- Utilisé par les chercheurs lorsqu'ils veulent faire des pré-tests afin de s'assurer que leurs instruments mesurent les bonnes choses;
- Aucune façon de savoir si les données recueillies sont représentatives, mais ce n'est pas ce le but.

8b. Échantillon volontaire

- Échantillon de gens qui sont volontaire pour répondre à un questionnaire, n'ont pas fait l'objet d'une sélection aléatoire, ont répondu à une petite annonce, ont fait un sondage en ligne;
- Utilisée en sociologie, psychologie, médecine. En psycho et médecine les chercheurs ne peuvent sélectionner de manière aléatoire les gens et les inciter à participer à une recherche médicale. Ils doivent recourir à la méthode volontaire.

Pour info... en psychologie,

- Les échantillons « expérimental » et de « contrôle » sont utilisés lorsque l'on fait une étude comparative,
- Et on parle d'échantillon « placebo » dans le cas d'individus sur lesquels on ne tente aucune expérience mais qui sont persuadés de faire partie de l'expérimentation.

Taille de l'échantillon

- Dépend de la précision des résultats que l'on veut obtenir;
- Dépend aussi des contraintes en termes de temps et de coûts.

La précision des résultats

Fait référence à la marge d'erreur (s'applique seulement aux échantillons probabilistes):

- C'est la différence entre la « vraie » valeur de la caractéristique d'une population et la valeur estimée à partir d'un échantillon probabiliste;
- C'est l'erreur d'estimation qu'on est disposé à accepter ou contrôler;
- C'est un pourcentage de la valeur du paramètre étudié dont on fixe comme erreur d'estimation raisonnable

Taux de confiance	Coefficient de marge (t)
90 %	1,65
95 %	1,96
99 %	2,57

En général, la taille de l'échantillon et la précision dépendent:

1. Du degré d'homogénéité de la population
2. De la taille de l'échantillon en soi

Problèmes (biais) des échantillons

- Imperfections dans les échantillons
- Erreur dans l'observation
- Les non-réponses, absences, refus
 - Les non-réponses peuvent être éloquentes...
- Dans le cas des échantillons probabilistes:
 - Plus l'échantillon est grand, plus il sera précis. Mais...
 - Le plus important n'est pas la taille mais sa représentativité

Questions!

1. On désire faire une enquête sur les goûts musicaux de la population suisse. Pour cela on choisit au hasard 1000 numéros de téléphone dans l'ensemble des annuaires et on appelle les répondants le jour. On obtient 583 réponses. Est-ce représentatif?

2. Pour connaître les opinions politiques de la population d'une ville on envoie 5 enquêteurs interroger les gens à la sortie de 5 grands magasins. Ils doivent questionner les clients jusqu'à ce qu'ils réunissent chacun un échantillon de 200 réponses. Est-ce représentatif?

3. Pour se faire une idée du niveau de la classe, on prend les étudiants des deux premières rangées et on les questionne. Est-ce représentatif?

**Que sait-on de l'expérience existentielle des parents
des jeunes contrevenants?**

Qu'avez-vous pensé de cet extrait?

Problématique

L'objet de notre recherche est d'introduire un sujet qui, à notre connaissance, n'a jamais été directement étudié. Il s'agit des impacts négatifs vécus par les parents des jeunes contrevenants(es). En effet, dans la littérature d'inspiration psychologique et sociologique, les parents des jeunes contrevenants sont généralement étudiés dans une optique de causalité. La question clef est essentiellement celle-ci : quelles sont les caractéristiques familiales qui auraient pu contribuer à la criminalité ou aux problèmes des enfants? En d'autres termes, quels sont les problèmes ou les anomalies présentées par ces parents?

...

Poursuivant les théories psychologiques et sociologiques visant à rétablir un certain équilibre dans la question de causalité en ce qui a trait aux comportements de l'enfant et de l'adolescent, nous avons émis l'hypothèse suivante : les parents des jeunes contrevenants se sentent affectés dans quatre domaines visés par notre étude : leur niveau de stress, de bonheur général, de fatigue et leur état général de santé.

Méthodologie

Notre recherche a été effectuée dans le cadre du fonctionnement quotidien du Centre de protection de l'enfance et de la jeunesse de Montréal. L'échantillon est constitué de jeunes âgés de 12 à 18 ans, recommandés au directeur pour évaluation-orientation, en vertu de la Loi sur les jeunes contrevenants. Les adolescents arrêtés et confiés au directeur sont rencontrés en compagnie d'au moins un de leurs parents par un délégué du Centre. Le but de cette rencontre est de procéder à une évaluation de la situation du jeune, de ses besoins spécifiques, de l'aspect délictuel et du risque de récidive qu'il présente.

... Nous avons demandé la collaboration de plusieurs délégués à la jeunesse afin qu'ils distribuent aux parents, après l'entrevue, un questionnaire de quatre pages. Celui-ci était accompagné d'une brève lettre expliquant le but de l'étude, assurant la confidentialité aux répondants et d'une enveloppe-réponse affranchie. Chaque questionnaire était numéroté et ce numéro, pour les besoins de la recherche, était inscrit sur la photocopie de la fiche d'information du jeune.

Le questionnaire se devait d'être bref et fort simple et ce pour trois raisons. Dans un premier temps, il fallait motiver les parents à le remplir. Deuxièmement, nous ne voulions pas trop en demander à ces parents qui sont déjà fort sollicités. Troisièmement, nous savions dès le départ qu'une large proportion des parents des jeunes contrevenants n'avait pas un degré de scolarité élevé. Ces considérations d'ordre pratique requéraient des questions à choix multiples...

132 questionnaires ont été distribués aux parents par les intervenants entre juin 1992 et octobre 1993. L'échantillon final inclut 116 questionnaires (ou 116 jeunes)... Notre échantillon est constitué de 78% de garçons et 22% de filles, respectivement âgés en moyenne de 16 et 14,9 ans lors de l'entrevue.

Plusieurs problèmes !!

1. Problème avec l'échantillon, qui ne correspond pas à la population étudiée :
 - 132 questionnaires ont été distribués aux parents par les intervenants entre juin 1992 et octobre 1993. L'échantillon final inclut 116 questionnaires (ou 116 jeunes). Notre échantillon est constitué de 78% de garçons et 22% de filles, respectivement âgés en moyenne de 16 et 14,9 ans lors de l'entrevue.

Un autre extrait qui concernait l'échantillon :

- Les familles étaient en majorité «fonctionnelles », en ce sens qu'elles comprenaient relativement peu de cas présentant des problèmes multiples. Cela peut en partie s'expliquer par le fait que les intervenants sociaux n'ont pas jugé pertinent de donner le questionnaire à des parents hostiles ou si marginaux que cette demande de collaboration aurait pu être mal reçue. De plus, l'échantillon excluait les parents qui ne maîtrisaient pas suffisamment la langue française ou qui étaient illettrés. Cette présélection, nécessaire à ce stade d'une recherche amenant une nouvelle perspective, nous laisse un échantillonnage excluant la plupart des parents qui ont eux-mêmes des problèmes sérieux et multiples, ainsi qu'un bon nombre de parents récemment immigrés.

2. Choix de méthode empreint de préjugés et d'éléments discriminants :

- Le questionnaire se devait d'être bref et fort simple et ce pour trois raisons. Dans un premier temps, il fallait motiver des parents de délinquants à le remplir. Deuxièmement, nous ne voulions pas trop en demander à ces parents qui sont déjà fort sollicités. Troisièmement, nous savions dès le départ qu'une large proportion des parents des jeunes contrevenants n'avait pas un degré de scolarité élevé. Ces considérations d'ordre pratique requéraient des questions à choix multiples.

Et la problématique?

Il me semble qu'elle ne concorde pas tout à fait avec la méthode empruntée.

1. Une expérience existentielle se mesure généralement plus adéquatement avec une technique qualitative. Une expérience existentielle ça se situe entre une trajectoire et un récit de vie. Il aurait fallu selon moi prendre un autre terme que « expérience existentielle ».
2. Dans la problématique les liens de causalité entre les caractéristiques familiales et la criminalité des enfants sont soulevés, mais aucune articulation n'est faite sur ce lien de causalité dans leur hypothèse.
3. Leur cadre analytique est vague, entre des théories psychologiques et des théories sociologiques... On ne sait pas quel genre d'analyse ils mèneront. Est-ce qu'ils feront des analyses régressives, analyses factorielles, analyses de correspondance...