

Tirage en grappes (plan SIGR)

Pierre Duchesne

August 1, 2017

- ▶ Jusqu'à maintenant, on supposait que la population U était fournie et que l'on pouvait lister les éléments de la population.
- ▶ Les plans de grappes sont particulièrement utiles lorsque l'on ne peut pas lister toutes les unités de U de 1 à N .
- ▶ Quand il y a absence de base de sondage, les plans de grappes sont très utiles.

Combien de ménages de Longueuil possèdent une piscine?

- ▶ Option 1: On pourrait prendre un échantillon de résidents selon un tirage SI, disons $n = 500$. Ceci peut entraîner des coûts au niveau du déplacement dans Longueuil.
- ▶ Option 2: On pourrait diviser Longueuil en pâtés de maison d'environ 20 ménages et prendre un échantillon de pâtés de maison. Une fois dans un quartier, on devrait sonder les maisons du coin. On pourrait donc lister tous les pâtés de maison dans un échantillon de 25 pâtés de maison.
- ▶ L'option 2 est un exemple de plan de grappes.

- ▶ Unités d'échantillonnage primaires: ce sont dans notre exemple les pâtés de maisons (en anglais, on trouve l'abréviation *psu*). Ce sont les unités d'échantillonnages (unités qui sont choisies par le plan de sondage).
- ▶ Unités d'échantillonnage secondaires: ce sont les ménages (*ssu*). Ici ce sont les unités d'analyse, que l'on veut mesurer.

Autres exemples

	grappes	unités d'analyse	variable d'intérêt (y)
1.	pâtés de maison,	ménages,	dépenses alimentaires;
2.	zones dans une forêt,	arbres,	volume du bois;
3.	ménages,	membres adultes,	posséder un travail.

Les situations 1. et 2. sont des exemples de sondages dits aréolaires.

Supposons que l'on dispose de notre échantillon de grappes. Pour les grappes sélectionnées, on peut considérer les choix suivants:

- ▶ On décide d'observer toutes les unités dans les grappes. C'est un exemple de plan de grappes à un degré.
- ▶ On liste toutes les unités dans chaque grappe choisie et dans cette liste procéder à un sélection selon un plan de sondage. C'est un exemple d'échantillonnage à deux degrés.

Unités d'analyse et d'échantillonnage dans un plan à deux degrés

- ▶ Supposons qu'un plan à deux degrés est mis en oeuvre.
- ▶ On doit choisir un plan de sondage dans chaque grappe sélectionnée.
- ▶ Dans un tel plan, les unités d'analyse sont également des unités d'échantillonnage au second degré.

Avantages du plan de grappes

- ▶ Même si l'on ne dispose pas d'une base de sondage, on peut tout de même s'assurer que toutes les unités dans la population U possède une probabilité d'inclusion strictement supérieure à zéro; c'est-à-dire $\pi_k > 0, \forall k \in U$.
- ▶ Souvent moins dispendieux à organiser. Sonder la population canadienne par un plan SI pourrait engendrer de nombreux déplacements. Un plan de grappes pourrait permettre de contrôler cet aspect.

Inconvénients du plan de grappes

- ▶ Habituellement, pour une taille échantillonnale similaire, le plan de grappes donne moins de précision qu'un tirage aléatoire simple sans remise (plan SI).
- ▶ Par exemple, dans l'exemple des ménages de Longueuil, les ménages d'une même grappe ont tendance à se ressembler (valeurs des maisons, quartiers de professionnels versus ouvriers, jeunes versus personnes plus âgées, etc.).
- ▶ Une grappe n'est pas toujours représentative de la population.

Distinctions entre strates et grappes

Les strates sont des sous-populations telles que:

- ▶ On peut lister, strate par strate, les unités d'analyse dans les strates.
- ▶ On prend un échantillon dans chaque strate, $h = 1, \dots, H$.

Les grappes sont des sous-populations telles que:

- ▶ On ignore, avant le tirage, l'identité des unités d'analyse dans une grappe.
- ▶ On peut cependant lister les grappes (base de sondage avec les unités étant les grappes).
- ▶ Dans les grappes sélectionnées, on observe ou bien toutes les unités d'analyse (tirage de grappes), ou bien un sous-échantillon (tirage à deux ou plusieurs degrés).
- ▶ La taille de la population N pourrait être inconnue.