

Sondages stratifiés

Pierre Duchesne

August 1, 2017

Sondages stratifiés: quelques motivations

- ▶ Dans un plan aléatoire simple (plan SI): La sélection de l'échantillon se fait totalement au hasard, et ne fait aucune usage de renseignements que l'on pourrait posséder sur les membres de la population.
- ▶ Dans un plan stratifié: On utilise de telles données pour accroître l'efficacité (on vise la réduction de la variance) de la démarche.
- ▶ On a besoin d'une partition de la population en sous-groupes *relativement homogènes* (le strates). On procède à la sélection d'échantillons *indépendants* dans chaque strate.

Échantillonnage dans chaque strate de façon indépendante

- ▶ Il faut préciser comment chaque échantillon sera sélectionné dans chaque strate.
- ▶ On pourrait adopter le plan BE dans chaque strate, le plan SI dans chaque strate, ou encore le plan SY dans chaque strate.
- ▶ On pourrait combiner des plans d'échantillonnage.
- ▶ On va porter une attention particulière au plan SI dans chaque strate: plan STSI.

Étude de marché: enquête portant sur le nombre d'heures hebdomadaire sur le réseau social Instagram

- ▶ Un pourcentage non-négligeable d'internautes utilise les applications mobiles des réseaux sociaux pour rechercher des produits à acheter.
- ▶ L'échantillon devrait représenter adéquatement les hommes et les femmes.
- ▶ L'utilisation d'Instagram varie beaucoup selon l'âge. Il serait donc bon d'avoir un échantillon qui représente bien les tranches d'âge.
- ▶ Si on a ces renseignements dans la base de sondage (sexe, âge), on peut les utiliser avec un plan STSI.

- ▶ Elles peuvent être qualitatives ou quantitatives.
- ▶ Géographiques: région, province, zone rurale, zone urbaine.
- ▶ Administratives: centre régionaux.
- ▶ Socio-démographiques: groupes d'âge, sexe, occupation.
- ▶ Taille: mesure de taille quantitatives

Avantages de l'échantillonnage stratifié: représentativité

- ▶ Échantillonnage qui permet de réduire le nombre d'échantillons possibles à ceux qui représentent mieux la population.
- ▶ Il est espéré que les échantillons possibles sont des 'modèles réduits' de la population.
- ▶ Représentativité est améliorée dans le sens suivant: comme il y a élimination (réduction) des échantillons qui ne sont pas des copies fidèles de la population, il y a moins d'échantillons qui vont produire des estimations éloignées des vraies valeurs des paramètres.
- ▶ L'idée principale est que les échantillons possibles devraient bien représenter tous les éléments de la population.

Autres avantages de l'échantillonnage stratifié

- ▶ Précision: Plus les strates sont homogènes, plus les estimations sont précises. Il est possible d'élaborer une théorie qui formalise ceci.
- ▶ Estimation distinctes dans les strates: On peut obtenir des estimations séparées pour chaque strate. Ainsi dans l'exemple d'Instagram il peut être pertinent d'avoir de l'information sur une tranche d'âge précise (exemple: les 16-24 ans).
- ▶ Un avantage important est qu'en plus d'avoir des estimations séparées, on a souvent une précision plus élevée (souvent beaucoup plus) des estimations au niveau de la population.

- ▶ **Enquêtes auprès des ménages:**
 1. Estimation par province: tranche de revenu; profession: groupe d'âges, etc.
- ▶ **Enquêtes auprès des entreprises**
 1. Estimation par activité économique; Statistique Canada a élaboré une classification en 11 divisions.
 2. Secteur géographique; limites de municipalités.