

Échantillonnage stratifié

Lorsqu'on utilise l'*échantillonnage stratifié*, on divise la population en groupes homogènes (appelés strates), qui sont **mutuellement exclusifs**, puis on sélectionne à partir de chaque strate des échantillons indépendants. On peut utiliser n'importe quelle des méthodes d'échantillonnage mentionnées dans la présente section (et il en existe d'autres) pour sélectionner l'échantillon à l'intérieur de chaque strate. La méthode d'échantillonnage peut varier d'une strate à une autre. Lorsqu'on utilise l'échantillonnage aléatoire simple pour sélectionner l'échantillon à l'intérieur de chaque strate, on appelle le plan d'échantillonnage un plan d'échantillonnage aléatoire simple stratifié. On peut stratifier avant l'échantillonnage une population au moyen de toute variable dont on dispose pour la totalité des unités incluses dans la base de sondage (comme l'âge, le sexe, la province de résidence, le revenu, etc.)

Pourquoi doit-on créer des strates? Pour bien des raisons, la principale étant que leur création peut rendre la stratégie d'échantillonnage plus efficace. Nous avons mentionné précédemment que vous aviez besoin d'un échantillon plus grand pour obtenir une estimation plus exacte d'une caractéristique qui varie beaucoup d'une unité à l'autre. Si chaque personne incluse dans une population, par exemple, avait le même salaire, il suffirait alors d'un échantillon d'une seule unité pour obtenir une estimation précise du salaire moyen des membres de cette population.

C'est l'idée qui sous-tend le gain d'efficacité qu'on réalise grâce à la stratification. Si vous créez des strates à l'intérieur desquelles des unités auraient des caractéristiques qui seraient similaires (comme le revenu) et qui différeraient considérablement de celles d'unités incluses dans d'autres strates (comme la profession et le type de logement), vous n'auriez alors besoin que d'un petit échantillon tiré de chaque strate afin d'obtenir une estimation précise du revenu total pour la strate en question. Vous pourriez ensuite combiner ces estimations afin d'obtenir une estimation précise du revenu total de l'ensemble de la population. Si vous deviez utiliser un échantillonnage aléatoire simple de la population entière sans effectuer de stratification, il vous faudrait un échantillon plus grand que la totalité de tous les échantillons de strate afin d'obtenir pour le revenu total une estimation du même degré de précision.

L'échantillonnage stratifié nous assure d'obtenir une taille d'échantillon suffisante pour des sous-groupes de la population à laquelle nous nous intéressons. Étant donné que chaque strate devient une population indépendante lorsque vous stratifiez une population, vous devrez déterminer pour chaque strate la taille de l'échantillon.

Exemple n° 6 : Supposez que vous voulez estimer combien d'élèves des écoles secondaires ont un emploi à temps partiel, et ce, tant au niveau national qu'à celui de chaque wilaya.

Si vous deviez sélectionner un échantillon aléatoire simple de 25 000 personnes à partir d'une liste de tous les élèves des écoles secondaires de notre pays (en supposant que vous disposiez d'une telle liste pour effectuer cette sélection), Le fait de stratifier votre liste par wilaya, en supposant encore une fois que vous disposeriez de cette information, puis de sélectionner une taille d'échantillon pour chacune des wilaya vous permettrait de déterminer la taille d'échantillon exacte qu'il vous faudrait.

En a 48 wilaya, en sélectionne pour chaque wilaya le nombre d'élève des écoles secondaire qui feront partis de notre échantillon pour chaque wilaya (puisque en a les listes des élèves pour chaque wilaya) par le biais de la méthode de l'échantillon aléatoire simple.

Supposant qu'on a 1200000 élèves des écoles secondaires sur le territoire national qui sont répartie uniformément dans les 48 wilaya soit 25000 élèves par wilaya, en voudrait sélectionné un échantillon de 60000 qui représentera tout les élèves du pays, en divise le nombre de l'échantillon 60000 sur 48 wilaya, en obtiendras 1250 élèves par wilaya. Par le biais de la liste des élèves et par le biais de la liste de la table de l'échantillon aléatoire, en sélectionne les élèves qui feront partie de notre échantillon soit 1250 élèves pour chaque strate (wilaya).