



L'Association internationale des statisticiens d'enquêtes (IASS en anglais, <http://isi-iass.org/home/>) et l'Institut de statistique de l'Université de Neuchâtel, Suisse organisent le séminaire en distanciel:

### **La méthode généralisée du partage des poids et ses applications**

**Date : 26.02.2021**

**Heure : 14.00-15.00 (CET)**

#### **Intervenants:**

- **Pierre Lavallée** (statisticien d'enquête, Canada)
- **Guillaume Chauvet** (ENSAI, IRMAR, France)

**Modératrice: Alina Matei** (Université de Neuchâtel, Suisse)

Pour assister au séminaire, veuillez vous enregistrer à l'adresse

<https://attendee.gotowebinar.com/register/5878493150998947087>

Après enregistrement, vous recevrez un courriel de confirmation avec l'information nécessaire pour y assister. Le format des interventions sera de 25 minutes, suivi de questions.

#### **Résumés des interventions**

- 1) Pierre Lavallée : Le sondage indirect – un survol

Pour tirer les échantillons nécessaires aux sondages sociaux ou économiques, il arrive qu'on ne dispose pas directement d'une base de sondage pour la population cible, mais plutôt d'une base de sondage d'une autre population reliée d'une certaine façon à la population cible. On peut alors imaginer le tirage d'un échantillon de la base disponible afin de produire une estimation pour la population cible en se servant de la correspondance (liens) existante entre les deux populations. C'est ce qu'on peut désigner par *sondage indirect*. Lorsque les liens entre les deux populations ne sont pas bijectifs, l'estimation d'un total (ou d'une moyenne) peut constituer un défi de taille. Afin de résoudre ce problème, on a développé la *méthode généralisée du partage des poids* (MGPP) qui permet d'obtenir un poids d'estimation pour chaque unité enquêtée de la population cible. La présente communication se veut un survol entourant les différents développements effectués par l'auteur autour du sondage indirect et de la MGPP.

- 2) Guillaume Chauvet : Généralisation de la méthode du partage des poids à une population échantillonnée continue - Application à l'Inventaire National Forestier (France)

Dans une enquête, on peut distinguer plusieurs types d'unités statistiques. Les unités d'échantillonnage sont celles pour lesquelles on dispose d'une base de sondage, et qui peuvent donc être tirées selon un plan de sondage. Les unités d'observation sont celles qui sont effectivement enquêtées sur le terrain, et

qui constituent la population d'intérêt. Dans les cas les plus simples, ces deux types d'unités coïncident. C'est par exemple le cas si on sélectionne un échantillon d'individus qui sont ensuite enquêtés. Dans des cas plus complexes, ces deux types d'unités sont distincts. C'est par exemple le cas si on tire un échantillon d'individus, et qu'on enquête ensuite les ménages associés à ces individus.

Dans le cas où unité d'échantillonnage et unité d'observation diffèrent, l'inférence sur la population observée est plus complexe. Pour résoudre ce problème, Deville et Lavallée (2006) ont proposé la méthode de partage des poids, basée sur un principe de dualité entre population échantillonnée et population observée. Une variable d'intérêt est ainsi exprimée sous la forme d'une variable synthétique dans la population échantillonnée, ce qui permet d'utiliser directement les propriétés du plan de sondage pour réaliser l'inférence sur la population observée. Cette méthode conduit également à partager les poids des unités échantillonnées entre les unités observées, d'où son nom.

Deville et Lavallée (2006) ont considéré le cas où les deux populations étaient de nature discrète (cas discret-discret). Cela recouvre un grand nombre des situations possibles, mais pas toutes. Dans le cas de l'Inventaire National Forestier français, la population échantillonnée est de nature continue (le territoire français), et on tire sur ce territoire un échantillon de points. Autour de ces points sont tracés des placettes circulaires, et les arbres situés dans ces placettes sont enquêtés. La population observée (les arbres), bien que très grande, est de nature discrète.

Dans ce travail, nous étendons la méthode de partage des poids au cas d'une population échantillonnée continue et d'une population observée discrète (cas continu-discret). Nous montrons que la méthode se généralise naturellement dans ce cas, en illustrant notre approche par des exemples. Il s'agit d'un travail en cours, en collaboration avec Olivier Bouriaud (Laboratoire d'Inventaire Forestier, France) et Philippe Brion (IRMAR, France).

## **Biographies des intervenants**

**Pierre Lavallée** a étudié en mathématiques et informatique (B.Sc) à l'Université d'Ottawa où il s'est spécialisé en statistique. En 1985, il a commencé à travailler comme méthodologiste d'enquête à Statistique Canada. Tout au long de sa carrière, il a travaillé dans les enquêtes économiques, les enquêtes agricoles, ainsi que les enquêtes sociales. Après une maîtrise en statistique (M.Sc.) à l'Université Carleton, il a complété un doctorat (Ph.D.) avec Jean-Jacques Droesbeke et Jean-Claude Deville à l'Université Libre de Bruxelles où il a produit une thèse sur le sondage indirect. Cette thèse a fait l'objet d'un livre en français « Le Sondage Indirect, ou la Méthode Généralisée du Partage des Poids » (Éditions Ellipses) et en anglais « Indirect Sampling » (Springer). Depuis mai 2017, il est à la retraite, tout en continuant d'être actif au sein de la statistique d'enquête. Il donne régulièrement des cours et conférences sur différents sujets comme les enquêtes longitudinales (panels), le sondage indirect, les populations rares et l'utilisation de données massives dans un contexte probabiliste.

**Guillaume Chauvet** est docteur en Statistiques et habilité à diriger des recherches. Il est enseignant-chercheur à l'École Nationale de la Statistique et de l'Analyse de l'Information (ENSAI, France) et membre de l'Institut de recherche mathématique de Rennes (IRMAR, UMR-6625). Il s'intéresse aux multiples aspects de la Statistique d'Enquête, incluant les méthodes d'échantillonnage, le traitement de la non-réponse, l'estimation de précision, le traitement de données longitudinales, et l'analyse de données d'enquête. Il est également membre du fan-club des sweats à capuche.