# COORDINATION NÉGATIVE D'ÉCHANTILLONS : UN SURVOL DE QUELQUES MÉTHODES

Pierre-Olivier Julien[1], Christian Gagné[2] et Carlos A. Leon[3]

## RÉSUMÉ

Une quantité considérable de données est produite par les agences statistiques. Ceci ajoute une charge de travail parfois importante aux répondants. Plusieurs méthodes de coordination d'échantillons tirés à même une base de sondage unique furent développées pour réduire leur fardeau de réponse. On échafaude présentement à Statistique Canada un projet central qui couvrira plus de 100 enquêtes basées sur le Registre des entreprises. La coordination de tous ces échantillons pourrait être possible à partir d'un outil informatique commun. Cet article se veut un survol de certaines méthodes pouvant être implantées dans un tel outil.

MOTS CLÉS : Coordination négative, coordination d'échantillons, nombres aléatoires, ÉASS, fardeau de réponse, lissage.

## ABSTRACT

A considerable amount of data is produced by statistical agencies, which can add a significant workload on respondents. To reduce respondent's burden, a number of methods for sample coordination of surveys based on a common frame have been developed. Statistics Canada is currently putting in place a central project consisting of more than 100 surveys based on its Business Register. Coordination of these samples would be possible from a common informatics tool. We present here an overview of some sample coordination methods which could be implemented in such a tool.

KEY WORDS: Negative Coordination; Sampling Coordination; Random Numbers; SSRS; Response Burden; Smoothing.

## 1. INTRODUCTION

### 1.1 Description du problème

Plusieurs enquêtes auprès des entreprises de Statistique Canada sont tirées à même le Registre des entreprises (RE), une base de sondage unique contenant toutes les entreprises enregistrées au pays et qui est mise à jour à partir d'enquêtes et de données fiscales. Chacune est normalement dirigée par une équipe s'occupant de la méthodologie de l'enquête. Les plans de sondages utilisés pour mener celles-ci sont normalement basés sur des échantillons aléatoires simples stratifiés (ÉASS). Les plans de sondage sont normalement prédéterminés par les équipes d'enquête respectives. On échafaude présentement un projet central qui couvrira plus de 100 enquêtes basées sur le RE. Il serait intéressant de pouvoir coordonner ces nombreux échantillons à l'aide d'un outil informatique commun. Un système généralisé de sélection d'échantillons existe à Statistique Canada, mais aucune méthode de coordination d'échantillons n'y a été implantée. Puisque celui-ci est présentement remanié, on en profite pour regarder s'il serait possible d'ajouter une telle composante dans le système. On voudrait principalement que la méthode de coordination d'échantillons contrôle le fardeau cumulé par les entreprises répondantes à travers les multiples ÉASS tirés à partir du RE. Cette façon de faire se nomme : coordination négative d'échantillons.

Plusieurs méthodes ont été envisagées pour répondre à nos besoins. Le présent document en décrit quelques-unes qui ont été considérées et résume la théorie mathématique derrière elles. On y présente aussi les résultats obtenus suite à une simulation comparant les méthodes. Une discussion sur ces résultats, ainsi qu'un résumé des avantages et inconvénients de chacune y sont donnés. Une conclusion récapitulant les faits saillants et une brève description de la suite du projet d'implantation d'une telle composante terminent le document.

[1] Pierre-Olivier Julien, Statistique Canada, Tunney's Pasture, Ottawa, ON K1A 0T6, pierre-olivier.julien@statcan.gc.ca
[2] Christian Gagné, Statistique Canada, Tunney's Pasture, Ottawa, ON K1A 0T6, christian.gagne@statcan.gc.ca
[3] Carlos A. Leon, Statistique Canada, Tunney's Pasture, Ottawa, ON K1A 0T6, carlos.leon@statcan.gc.ca

## 2. REVUE DE LITTÉRATURE

### 2.1 Quelques méthodes connues de coordination d'échantillons

Plusieurs méthodes de coordination d'échantillons ont été développées au cours des années. Certaines sont même utilisées par des agences statistiques dans le monde. Rivière (1999) en énumère quelques-unes. De même, Nedyalkova et al (2006) en énumèrent et en comparent plusieurs, dont Kish et Scott, Néerlandaise (*Dutch method*) et l'utilisation de la programmation linéaire (*LP*), pour n'en nommer que quelques-unes. Ces méthodes ne répondent toutefois pas à nos besoins, puisque soit elles deviennent compliquées à mettre en œuvre lorsqu'on parle d'une série de plusieurs dizaines d'enquêtes à coordonner (Kish et Scott, LP), soit les plans de sondage ne peuvent être prédéterminés (*Dutch method*). Par contre, Nedyalkova et al (2006) ont aussi traité et comparé une panoplie d'autres méthodes qui répondraient plus à nos attentes : la méthode de Cotton et Hesse (1992) et celles plutôt basées sur la notion de micro-strates qu'on définira plus tard. On réfèrera à ces dernières dans le texte comme étant des méthodes dites de Permutation de Nombres Aléatoires basée sur le Fardeau de réponse (PNAF) cumulé ou non (Rivière (1999)).

### 2.2 Théorie derrière les méthodes de type PNAF

On définit d'abord une série d'enquêtes $(s^1, \ldots, s^K)$ toutes tirées d'une base de sondage unique contenant $N$ unités, où chaque $s^k$ est de type ÉASS basé sur un plan de sondage préétabli. On définit aussi un vecteur de nombres aléatoires (VNA) comme étant un vecteur contenant $N$ éléments venant d'une loi $U(0,1)$ de telle sorte que le $k^e$ VNA $= \omega^k = (\omega^k_1, \ldots, \omega^k_N)$. Une façon simple de tirer un ÉASS $s^k$ d'une telle base de sondage est de sélectionner les $n^k_h$ premiers éléments de la strate $h$, et ce, en triant selon l'ordre croissant des nombres aléatoires de $\omega^k$.

Le but maintenant est d'ajouter une $K+1^e$ enquête à la série. Pour ce faire, on aura besoin d'un $\omega^{K+1}$ afin de sélectionner l'échantillon de type ÉASS. Pour obtenir celui-ci, on suggère d'utiliser les méthodes de type PNAF à partir du dernier VNA de la série : $\omega^K$.

L'idée générale derrière ces méthodes est d'appliquer un algorithme de tris permettant de **permuter** les éléments de $\omega^K$, c'est-à-dire d'obtenir une relation un-à-un entre les éléments composant les vecteurs avant et après l'application de la méthode. Cela revient à poser $f(\omega^K) = \omega^{K+1}$, où $f$ est une permutation de type PNAF. Ceci implique un fait très important : si $\omega^K$ provient d'une distribution statistique quelconque, alors $\omega^{K+1} = f(\omega^K)$ aura la même distribution (Rubin-Bleuer et Rivière (2011)). Ainsi, si on applique ces méthodes à partir d'un VNA de départ provenant d'une distribution $U(0,1)$, les VNA résultants pour chaque enquête proviendront aussi d'une $U(0,1)$. Par conséquent, on obtiendra aussi un ÉASS en appliquant la façon de faire décrite au premier paragraphe du sous-chapitre.

Par contre, bien qu'il faille tenir compte de cette caractéristique obligatoire, la méthode doit aussi limiter le fardeau de réponse cumulé dans le temps par chaque unité de la base de sondage. On nomme ceci : lisser le fardeau, c'est-à-dire qu'étant donné les plans de sondage prédéterminés, et donc le nombre connu d'unités à sélectionner à travers le temps, on tentera de répartir uniformément (lisser) à travers les unités le nombre de fois qu'elles seront sélectionnées (fardeau cumulé) dans la série d'enquêtes tirée à partir de la même base de sondage. Le tableau 1 ci-dessous montre un exemple de permutation de nombres aléatoires basée sur le fardeau de réponse cumulé (FRC) à l'intérieur d'une même strate.

**Tableau 1 – Exemple d'une permutation de nombres aléatoires basée sur le fardeau de réponse cumulé**

| État | | Unité A | Unité B | Unité C | Unité D | Unité E | Unité F |
|---|---|---|---|---|---|---|---|
| Avant le tri | FRC | 3 | 2 | 0 | 1 | 2 | 0 |
| | # aléatoire | 0.93 | 0.36 | 0.40 | 0.12 | 0.52 | 0.25 |
| Après le tri | FRC | 3 | 2 | 0 | 1 | 2 | 0 |
| | # aléatoire | 0.93 | 0.40 | 0.25 | 0.36 | 0.52 | 0.12 |

On définit ensuite la notion de micro-strates, qui est très importante pour les méthodes de type PNAF. On sous-entend par *micro-strates-(m, K)* l'intersection de **toutes** les strates des enquêtes comprises entre la $m^e$ et la $K^e$ enquête de la série. Le tableau 2 ci-dessous montre un exemple de création de *micro-strates-(m, K)*.

**Tableau 2 – Exemple de création de micro-strates**

| Unité | Strates 1 | Strates 2 | Strates 3 | Micro-strates-(1, 3) | Micro-strates-(2, 3) |
|---|---|---|---|---|---|
| A | 11 | 21 | X | 11-21-X | 21-X |
| B | 11 | 21 | X | 11-21-X | 21-X |
| C | 11 | 22 | Y | 11-22-Y | 22-Y |
| D | 12 | 22 | Y | 12-22-Y | 22-Y |
| E | 12 | 23 | Y | 12-23-Y | 23-Y |
| F | 12 | 23 | Y | 12-23-Y | 23-Y |

On notera que les *micro-strates-(3, 3)* sont, par définition, les strates de la 3$^{e}$ enquête.

Finalement, on définit le *FRC-(m, K)* comme étant la somme des fardeaux de réponse assignés aux enquêtes (FRAE) de la série pour lesquelles l'unité a été sélectionnée. Un FRAE est défini comme un « coût » associé au fait de compléter un questionnaire pour l'enquête. Il est fixé par l'utilisateur. Si l'unité n'est pas sélectionnée, le FRAE est de zéro. Habituellement, lorsqu'on considère les enquêtes comme ayant des fardeaux de réponse semblables, la valeur 1 est utilisée. En reprenant l'exemple pris pour le tableau 2, le tableau 3 ci-dessous indique comment calculer le *FRC-(m, K)* à partir de FRAE préfixés par l'usager, soit 2 pour la première enquête et 1 pour les deux autres.

**Tableau 3 – Exemple de calcul du fardeau de réponse cumulé**

| Unité | Éch. 1 | Éch. 2 | Éch. 3 | FRAE 1 | FRAE 2 | FRAE 3 | FRC-(1, 3) |
|---|---|---|---|---|---|---|---|
| A | 1 | 0 | 1 | 2 | 0 | 1 | 3 |
| B | 0 | 1 | 1 | 0 | 1 | 1 | 2 |
| C | 0 | 1 | 1 | 0 | 1 | 1 | 2 |
| D | 1 | 0 | 1 | 2 | 0 | 1 | 3 |
| E | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| F | 1 | 0 | 0 | 2 | 0 | 0 | 2 |

Afin d'obtenir un VNA pour pouvoir tirer un ÉASS pour une *K+1*$^{e}$ enquête ajoutée à la série en contenant déjà *K*, on se doit d'appliquer ce qu'on appelle ici une *permutation-(m, K)*. Cette dernière est définie comme suit : étant donné une série de *K* enquêtes avec des plans de sondage de type ÉASS, une *permutation-(m, K)* est simplement une permutation des éléments de $\omega^{K}$ à l'intérieur des *micro-strates-(m, K)* de telle sorte que les nombres aléatoires de $\omega^{K}$ sont triés en ordre croissant de *FRC-(m, K)*. Le tableau 4 résume les notions décrites dans cette sous-section en complétant l'exemple.

**Tableau 4 – Exemple de création d'une *permutation-(m, K)***

| Unité | Micro-strates (1, 3) | FRC-(1, 3) | Micro-strates (2, 3) | FRC-(2, 3) |
|---|---|---|---|---|
| A | 11-21-X | 3 | 21-X | 1 |
| B | 11-21-X | 2 | 21-X | 2 |
| C | 11-22-Y | 2 | 22-Y | 2 |
| D | 12-22-Y | 3 | 22-Y | 1 |
| E | 12-23-Y | 1 | 23-Y | 1 |
| F | 12-23-Y | 2 | 23-Y | 0 |

### 2.3 Les méthodes de type PNAF considérées

Les méthodes qui ont été analysées et comparées sont : le Micro-strates de base (*Microstrates*), le global négatif (*Global*) et celle de Cotton et Hesse (*C&H*). Les deux premières étant décrites dans Rivière (1999) et commentées dans Rubin-Bleuer et Rivière (2011) et la troisième dans Cotton et Hesse (1992). L'idée derrière le Micro-strates de base est de lisser le fardeau à travers toutes les enquêtes de la série. Cette méthode crée en effet une relation un-à-un entre les éléments des

VNA selon Rubin-Bleuer et Rivière (2011). Par contre, l'application successive de cette méthode amène un problème majeur : les micro-strates sont toujours de plus en plus petites et la méthode devient par conséquent inefficace pour bien lisser le FRC. Le tableau 2 montrait bien cela puisqu'on passait de deux *micro-strates-(3, 3)* à trois *micro-strates-(2,3)* à quatre *micro-strates-(1,3)*. C'est pourquoi Rivière (1999) proposait donc d'effectuer trois tris successifs de sa méthode dans le global négatif : un tri avec la dernière enquête de la série, un tri lissant jusqu'à la première enquête de l'année calendrier et un dernier tri lissant jusqu'à deux années calendriers avant la nouvelle enquête ajoutée à la série. Finalement, Cotton et Hesse (1992) décrivent une méthode qui permute les nombres aléatoires sans regarder le fardeau cumulé par les unités. Seule la dernière enquête de la série est utilisée. Par contre, ils soulignent que le fait d'appliquer la méthode de façon consécutive lisse tout de même le fardeau cumulé par les unités. Le tableau 5 ci-dessous compare, en termes de *permutation-(m, K)*, les trois méthodes.

**Tableau 5 – Types de permutations basés sur le fardeau des méthodes PNAF comparées**

| Tris | ÉASS | Microstrata | Global | C&H |
|---|---|---|---|---|
| 1er | S.O. | (1, K) | (K, K) | (K, K) |
| 2e | S.O. | S.O. | ($K_{-1}$, K) | S.O. |
| 3e | S.O. | S.O. | ($K_{-2}$, K) | S.O. |

## 3. SIMULATIONS ET RÉSULTATS

### 3.1 Description de la base et des plans de sondage pour la simulation

Afin de comparer les méthodes, on a simulé une série de 10 enquêtes avec plans de sondage de type ÉASS. On a tenté de reproduire le mieux possible, mais à très petite échelle, une série classique et réaliste d'enquêtes auprès des entreprises tirées à même le RE. La taille de la base de sondage simulée était de 500 entreprises. À des fins de simplicité, on a supposé qu'aucune modification ne serait apportée à la base de sondage durant la série d'enquêtes, c'est-à-dire qu'aucune unité ne sera ajoutée ou enlevée de celle-ci. De plus, toujours afin de simplifier l'étude, les FRAE furent fixés à 1 pour toutes les enquêtes. Aussi, pour tester la méthode Global, on a supposé que la sixième enquête était la première de l'année courante. Ainsi, lorsque les enquêtes 7 à 10 seront ajoutées à la série pour simuler cette méthode, on fixera $K_{-1} = 6$. Le tableau 6 ci-dessous indique les types de permutations basées sur le fardeau utilisés pour la simulation.

**Tableau 6 – Types de permutations basées sur le fardeau des méthodes PNAF comparées (simulation)**

| Tris | ÉASS | Microstrata | Global (K<=6) | Global (K>6) | C&H |
|---|---|---|---|---|---|
| 1er | S.O. | (1, K) | (K, K) | (K, K) | (K, K) |
| 2e | S.O. | S.O. | (1, K) | (6, K) | S.O. |
| 3e | S.O. | S.O. | S.O. | (1, K) | S.O. |

Enfin, toujours à des fins de simplicité, on a supposé que chaque enquête serait formée de quatre strates : une à tirage nul, deux à tirage partiel (petites et moyennes entreprises) et une à tirage complet. Afin de dériver les strates, une variable de revenu a été créée de façon aléatoire. Le revenu étant la variable auxiliaire la plus souvent utilisée pour séparer les entreprises en groupe de taille, qui forment les strates de l'enquête. Les enquêtes n'ayant pas toutes nécessairement la même stratification, on a ajouté un effet aléatoire pour jouer avec les tailles de strates. Comme ce ne sont pas toutes les enquêtes-entreprises qui prennent le revenu pour variable de stratification, on a présumé que les enquêtes 5 et 10 seraient stratifiées selon une autre variable quelconque, simulée elle aussi. Finalement, on a fixé les tailles d'échantillon pour les strates à tirage partiel arbitrairement et elles sont toutes différentes d'une enquête à l'autre.

### 3.2 Résultats de la simulation et tableaux comparatifs

Pour comparer les méthodes, on a appliqué 5 000 répétitions Monte-Carlo pour chaque méthode PNAF en prenant comme points de départ différents $\omega^{1}$ tirés d'une *U(0,1)*. À des fins de comparaison, on a aussi simulé 10 ÉASS

indépendants pour notre série d'enquêtes, toujours à partir de 5 000 répétitions Monte-Carlo, et ce, pour chaque enquête séparément. À la fin de chaque simulation, le FRC de chacune des 500 entreprises a été calculé en sommant le nombre de fois qu'une unité a été sélectionnée. Les FRC varient entre 0, si l'entreprise n'a jamais été sélectionnée durant la série d'enquêtes, et 10, si l'entreprise a été sélectionnée pour chacune d'elles. Le tableau 7 ci-dessous montre la distribution des FRC à travers les simulations. Notons que le tableau indique les proportions de FRC obtenues pour les 500 entreprises à travers les 5 000 simulations Monte-Carlo pour un total de 2 500 000 FRC calculés en tout.

**Tableau 7 – Distribution en pourcentage (%) du fardeau cumulé par les unités selon la méthode PNAF**

| FRC | ÉASS | Microstrates | Global | C&H |
|---|---|---|---|---|
| 0 | 21,09 | 11,42 | 9,06 | 9,04 |
| 1 | 30,51 | 40,24 | 40,09 | 40,11 |
| 2 | 21,22 | 26,31 | 30,27 | 30,27 |
| 3 | 12,08 | 10,01 | 9,82 | 9,84 |
| 4 | 7,06 | 4,99 | 3,96 | 3,94 |
| 5 | 4,19 | 3,61 | 4,13 | 4,13 |
| 6 | 2,30 | 2,08 | 1,94 | 1,94 |
| 7 | 1,04 | 0,74 | 0,42 | 0,42 |
| 8 | 0,29 | 0,29 | 0,10 | 0,10 |
| 9 | 0,23 | 0,33 | 0,20 | 0,20 |
| 10 | 0,00 | 0,01 | 0,00 | 0,00 |
| Total | 100,00 | 100,00 | 100,00 | 100,00 |

Il est bien de noter que seul le Microstrates de base a créé des FRC de 10 dans les simulations. Ceci indique bien que lorsque les micro-strates sont réduites à une seule unité, non seulement la méthode ne lisse plus bien le fardeau, mais elle en ajoute continuellement pour une unité ayant un nombre aléatoire petit prise dans une telle micro-strate. Le FRC moyen espéré étant de 1,81, on remarque que les méthodes Global et C&H tendent à avoir des FRC plus prêts de cette valeur comparativement aux deux autres méthodes. Ceci est un indicateur que les deux méthodes lissent mieux le fardeau cumulé. Le tableau 8 montre les écarts-types calculés pour les FRC obtenus durant les simulations. Ils confirment bien que les méthodes Global et C&H lissent mieux le fardeau cumulé que le Microstrates de base, et surtout que les ÉASS indépendants, car leur étendu autour de la moyenne est plus petit.

**Tableau 8 – Écarts-types du fardeau cumulé par les unités selon la méthode PNAF**

| ÉASS | Microstrates | Global | C&H |
|---|---|---|---|
| 1,002 | 0,705 | 0,412 | 0,412 |

### 3.3 Discussions en vue d'un ajout au système informatique généralisé

On remarque dans le tableau 7 que les résultats obtenus pour les méthodes Global et C&H sont très semblables. On suppose qu'il n'en serait pas le cas si on avait utilisé des FRAE différents, mais ceci devra faire partie d'une étude ultérieure. Il n'en demeure pas moins que la simulation a démontré que ces deux méthodes lissent mieux le FRC que le Microstrates de base et l'utilisation d'ÉASS indépendants. L'utilisation des FRAE est l'atout le plus intéressant des méthodes Microstrates de base et Global. Il faudrait éventuellement tester l'effet de ne pas tenir en compte des FRAE dans l'application de C&H sur les FRC. Par contre, puisque la valeur associée d'un FRAE est déterminée de façon subjective (temps, longueur de questionnaire, mode de collecte, questions délicates, etc.) on considère souvent, par simplicité, chaque enquête comme ayant le même « coût ». Ainsi, la méthode de C&H, étant beaucoup plus simple à intégrer à un système informatique puisqu'il suffit de ne garder en mémoire que le plan de sondage de la dernière enquête de la série, serait un outil plus envisageable à court terme.

Néanmoins, il faut souligner qu'il y a d'autres particularités que l'on doit regarder avant d'implanter un tel outil dans les systèmes. Par exemple, plusieurs plans d'enquête prévoient l'utilisation d'un chevauchement positif ou taux de rotation

avec une ou certaines enquêtes. Serait-il possible d'appliquer une coordination positive avec une enquête donnée tout en lissant le fardeau de réponse cumulé avec les autres enquêtes lorsqu'on ajoute une nouvelle enquête à la série? Ceci devra aussi faire partie d'analyses futures. Par contre, il semblerait que la méthode Global pourrait répondre à ces deux besoins. Sinon, il faudrait plutôt envisager de ne pas ajouter dans la série une enquête qui a été coordonnée positivement avec une autre afin de ne pas « briser » la série de coordination négative. On donne au tableau 9 quelques notes sur certaines caractéristiques importantes concernant les méthodes PNAF que l'on a testées pour le système. Une note de 1 représentant la meilleure et 4 la moins bonne performance pour cette caractéristique.

**Tableau 9 – Notes attribuées à certaines caractéristiques des méthodes PNAF comparées**

| Caractéristiques | ÉASS | Microstrates | Global | C&H |
|---|---|---|---|---|
| Simplicité d'implantation | 1 | 3 | 4 | 2 |
| Performance au lissage | 4 | 3 | 1 | 1 |
| Utilisation de FRAE | Non | Oui | Oui | Non |
| Désavantage principal | Pas de coordination | Problème des petites micro-strates | Beaucoup d'espace mémoire requise | Coordination positive doit être faite hors de la série |

## 4. Conclusion

Afin d'inclure une composante de coordination d'échantillons à notre système généralisé de sélection d'échantillons, plusieurs méthodes de coordination connues ont été prises en compte. Par contre, les méthodes de type PNAF s'avèrent les plus intéressantes dans notre cas, puisqu'on veut pouvoir coordonner des centaines d'enquêtes de type ÉASS tirées à même le RE. De plus, leurs plans de sondage respectifs étant pour la majorité préétablis, seules les méthodes de type PNAF pouvaient aussi respecter ce fait.

Trois méthodes de type PNAF ont été comparées à l'aide d'une simulation. Elles ont par la suite été comparées entre elles, ainsi qu'avec une séquence d'ÉASS indépendants, autant à partir des résultats de la simulation qu'à l'aide de leurs avantages et inconvénients respectifs. Plusieurs analyses sont encore à effectuer avant d'aller de l'avant avec l'implantation d'un tel outil; l'ajout de différents FRAE à la simulation, la coordination positive, le calendrier de la série d'enquêtes, les besoins des utilisateurs, pour n'en nommer que quelques-unes. Néanmoins, on peut souligner que la méthode C&H serait une méthode de type PNAF pouvant être mise en place rapidement dans nos systèmes. Du moins, un module de coordination négative lissant le FRC sur une longue période serait disponible et les enquêtes pour lesquelles une coordination positive est nécessaire seraient mises en dehors de la série d'enquêtes coordonnées négativement.

## RÉFÉRENCES


Cotton, F. et Hesse, C. (1992). « Co-ordinated Selection of Stratified Samples ». *Proceedings of Statistics Canada Symposium 92*, **92**:47-54.

Nedyalkova, D., Pea, J. et Tillé, Y. (2006) « A Review of Some Current Methods of Coordination of Stratified Samples. Introduction and Comparison of New Methods Based on Microstrata ». *Rapport interne de l'Université de Neuchâtel,* Site web : https://libra.unine.ch/Publications/Personne/T/Yves_Tille/P-6

Rivière, P. (1999). « Coordination of Samples: the Microstrata Methodology ». *13th International Roundtable on Business Survey Frames*, Paris: INSEE.

Rubin-Bleuer, S. et Rivière, P. (2011). « Random Permutations Method of Sampling Coordination ». *Rapport interne DRIS-2011-001-E*, Statistique Canada.