Techniques d'enquête, juin 1979
Vol. 5, n° 1
Statistique Canada, n° 12-001-X au catalogue

Sélection d'un échantillon de taille n selon un plan avec probabilité proportionnelle à la taille sans remise dans une population finie

G.H. Choudhry[1]

Résumé

Soit U = {1, 2, 3, ..., i, ..., N} une population finie de N unités indentifiables. Une « mesure de la taille » connue x_i est associée à l'unité i, i = 1, 2, …, N. L'auteur propose une méthode d'échantillonnage pour choisir une taille d'échantillon n (2 < n < N) dont la probabilité est proportionnelle à la taille et sans remise. De cette façon la probabilité d'inclusion est proportionnelle à la taille pour chaque unité de la population.

Mots-clés : Probabilité d'inclusion; plan de sélection; mesure de la taille.

1 Introduction

Yates et Grundy [4] ont envisagé une procédure de sélection qui permet de choisir n unités selon un plan avec probabilité proportionnelle à la taille sans remise (PPTSR) lorsqu'un premier tirage effectue la sélection d'une unité par probabilité proportionnelle à la taille, qu'un deuxième tirage permet de sélectionner une unité dans le reste des unités à l'aide du même procédé, et ainsi de suite. Cependant, dans cette procédure, la probabilité générale d'inclusion d'une unité dans l'échantillon n'est pas proportionnelle à sa taille. Fellegi [1] a proposé une méthode dans laquelle la probabilité de sélectionner une unité est proportionnelle à sa taille dans chacun des n tirages successifs. Pour ce faire, il faut déterminer n ensembles de probabilités de sélection appelés « probabilités de travail ». Ainsi, la probabilité d'inclusion est proportionnelle à la taille de chacune des n unités de la population. Cette méthode exige toutefois une évaluation fastidieuse des « probabilités de travail » à chaque tirage, sauf au premier.

Dans la méthode actuelle, le plan de Yates et Grundy [4] est appliqué à la sélection des premiers (n - 1) tirages, et un ensemble de « probabilités de travail » est établi pour la sélection d'une unité au

1. G.H. Choudhry, Division des méthodes d'enquête — recensement et enquêtes-ménages, Statistique Canada.

énième tirage, de sorte que la probabilité générale d'inclusion d'une unité dans l'échantillon devient proportionnelle à la taille de chaque unité de la population. Des résultats empiriques montrent que cette méthode est aussi efficace que celle de Fellegi [1].

2 Méthode d'échantillonnage

Définissez les « tailles normalisées » p_i comme étant proportionnelles à x_i; i = 1, 2,…, N de sorte que $\sum_{i=1}^{N} p_i = 1$, c.-à-d.

$$p_i = \frac{x_i}{\sum_{i=1}^{N} x_i} \qquad i = 1, 2, \ldots, N. \tag{2.1}$$

Soit Π_i la probabilité de la présence d'une unité i dans l'échantillon; nous pouvons ensuite démontrer que

$$\sum_{i=1}^{N} \Pi_i = n \quad . \tag{2.2}$$

La probabilité d'inclusion de Π_i doit être proportionnelle à p_i; i = 1, 2,…, N. Cette condition, ainsi que celles des formules (2.1) et (2.2), signifie que

$$\Pi_i = np_i \qquad i = 1, 2, \ldots, N \tag{2.3}$$

Au premier tirage, la sélection d'une unité se fait selon une probabilité proportionnelle à la taille, un deuxième tirage permet de sélectionner une unité dans le reste des unités à l'aide du même procédé, et ainsi de suite jusqu'au $(n - 1)^e$ tirage. La probabilité de sélectionner la $i_1{}^e$ unité lors du premier tirage correspond à p_{i_1}, la probabilité conditionnelle de sélectionner la $i_2{}^e$ unité lors du deuxième tirage (étant donné que la $i_1{}^e$ unité a déjà été sélectionnée lors du premier tirage) est équivalente à

$$\frac{p_{i_2}}{1 - p_{i_1}}$$

et ainsi de suite. La probabilité conditionnelle de sélectionner la i_{n-1} e unité lors du (n - 1)e tirage {étant donné que les unités i_1, i_2, …, i_{n-2} ont déjà été sélectionnées auparavant) est équivalente à

$$\frac{p_{i_{n-1}}}{1 - p_{i_1} - p_{i_2} \ldots - p_{i_{n-2}}} .$$

Soit q_i; i = 1, 2,..., N l'ensemble de « probabilités de travail » pour la sélection lors du énième tirage, la probabilité conditionnelle de sélectionner la i_ne unité lors du énième tirage (étant donné que i_1, i_2, …, i_{n-1} ont déjà été sélectionnés) est équivalente à

$$\frac{q_{i_n}}{1 - q_{i_1} - q_{i_2} \ldots - q_{i_{n-1}}} .$$

Ensuite, la probabilité générale, $\delta_i(k)$, de sélectionner la ie unité lors du ke tirage s'exprime ainsi

$$\delta_i(k) = \sum_{(k-1,i)} p_{i_1} \times \frac{p_{i_2}}{1-p_{i_1}} \times \frac{p_{i_3}}{1-p_{i_1}-p_{i_2}} \ldots \times \frac{p_{i_{k-1}}}{1-p_{i_1}-p_{i_2}-\ldots-p_{i_{k-2}}} \times$$

$$\frac{p_i}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}} \qquad k = 1, 2, \ldots, n-1 \qquad (2.4)$$

et

$$\delta_i(n) = \sum_{(n-1,i)} p_{i_1} \times \frac{p_{i_2}}{1-p_{i_1}} \times \frac{p_{i_3}}{1-p_{i_1}-p_{i_2}} \times \ldots \times \frac{p_{i_{n-1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{n-2}}} \times$$

$$\frac{q_i}{1-q_{i_1}-q_{i_2} \ldots -q_{i_{n-1}}} \qquad (2.5)$$

où $\sum_{(k-1, i)}$ (selon l'étude de Fellegi [1]) indique la totalisation de toutes les unités possibles (k - 1) des tuples ordonnés de (i_1, i_2, …, i_{k-1}) de façon à ce que i_1, i_2, …, i_{k-1} soient des nombres entiers différents de 1 à N, et qu'aucun d'entre eux ne soit équivalent à i. Ensuite, Π_i, la probabilité de la présence de la ie unité dans l'échantillon, s'obtient avec la formule suivante :

$$\Pi_i = \sum_{k=1}^{n} \delta_i(k)$$

$$= \sum_{k=1}^{n-1} \delta_i(k) + \delta_i(n)$$

où $\delta_i(k)$ pour $k = 1, 2, \ldots, n-1$ s'obtient avec la formule (2.4) et $\delta_i(n)$ s'obtient avec la formule (2.5). Dans l'expression de $\delta_i(n)$, il faut déterminer $q_i; i = 1, 2, \ldots, N$ pour que la condition $\Pi_i = np_i$; $i = 1, 2, \ldots, N$ soit satisfaite, c'est-à-dire que

$$np_i = \sum_{k=1}^{n-1} \delta_i(k) + \sum_{(n-1,i)} p_{i_1} \times \frac{p_{i_2}}{1-p_{i_1}} \times \ldots \times \frac{p_{i_{n-1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{n-2}}} \times$$

$$\frac{q_i}{1-q_{i_1}-q_{i_2}\ldots-q_{i_{n-1}}} \qquad i = 1, 2, \ldots, N$$

$$q_i = \frac{np_i - \sum_{k=1}^{n-1} \delta_i(k)}{\sum_{(n-1,i)} \frac{p_{i_1} p_{i_2} \ldots p_{i_{n-1}}}{(1-p_{i_1})(1-p_{i_2}) \ldots (1-p_{i_1}-p_{i_2}\ldots-p_{i_{n-2}})(1-q_{i_1}-q_{i_2}\ldots-q_{i_{n-1}})}}$$

$$i = 1, 2, \ldots N. \qquad (2.6)$$

L'ensemble de « probabilités de travail » q_i; $i = 1, 2, \ldots, N$ s'obtient en trouvant la solution de l'ensemble d'équations non linéaires simultanées obtenues avec la formule (2.6), grâce à une procédure itérative dans laquelle la valeur initiale de q_i peut correspondre à p_i; $i = 1, 2, \ldots, N$.

Pour $n = 2$, la méthode est la même que celle présentée par Fellegi [1], mais pour $n \geq 3$, nous pouvons rendre Π_i équivalent à np_i pour tous les i en évaluant seulement un ensemble de probabilités de travail au lieu

des (n - 1) ensembles de probabilités de travail prévus dans la méthode de Fellegi.

3 Calcul de Π_{ij}

La probabilité conjointe d'inclusion des unités i et j dans l'échantillon Π_{ij}, i = 1, 2, ..., N − 1; j = i + 1, ..., N peut se calculer comme suit :

Soit $\delta_{ij}(k, \ell)$ la probabilité d'avoir sélectionné l'unité i au k^e tirage et l'unité j, au ℓ^e tirage, où $k < \ell$. Ensuite, $\delta_{ij}(k, \ell)$ s'obtient avec la formule

$$\delta_{ij}(k,\ell) = \sum_{(\ell-2,i,j)} p_{i_1} \times \frac{p_{i_1}}{1-p_{i_1}} \times \frac{p_{i_3}}{1-p_{i_1}-p_{i_2}} \times \ldots \times \frac{p_{i_{k-1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-2}}} \times$$

$$\frac{p_i}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}} \times \frac{p_{i_{k+1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}-p_i} \times$$

$$\ldots \times \frac{p_{i_{\ell-1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}-p_i-p_{i_{k+1}}\ldots-p_{i_{\ell-2}}} \times$$

$$\ldots \times \frac{p_j}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}-p_i-p_{i_{k+1}}\ldots-p_{i_{\ell-1}}} .$$

$$k = 1, 2, \ldots, n-2$$
$$\ell = k+1, \ldots, n-1$$

et $\delta_{ij}(k, n)$ s'obtient avec la formule

$$\delta_{ij}(k,n) = \sum_{(n-2;i,j)} p_{i_1} \times \frac{p_{i_2}}{1-p_{i_1}} \times \frac{p_{i_3}}{1-p_{i_1}-p_{i_2}} \times \ldots \times \frac{p_{i_{k-1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-2}}} \times$$

$$\frac{p_i}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}} \times \frac{p_{i_{k+1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}-p_i} \times$$

$$\ldots \times \frac{p_{i_{n-1}}}{1-p_{i_1}-p_{i_2}\ldots-p_{i_{k-1}}-p_i-p_{i_{k+1}}\ldots-p_{i_{n-2}}} \times$$

$$\ldots \times \frac{q_j}{1-q_{i_1}-q_{i_2}\ldots-q_{i_{k-1}}-q_i-q_{i_{k+1}}\ldots-q_{i_{n-1}}}$$

$$k = 1, 2, \ldots n-1,$$

où $\sum_{(\ell-2;\, i,\, j)}$ indique la totalisation de $(\ell - 2)$ - tous les tuples ordonnés de $(i_1, i_2, \ldots, i_{k-1}, i_{k+1}, \ldots, i_{\ell-1})$, de sorte que $i_1, i_2, \ldots, i_{k-1}, i_{k+1}, \ldots, i_{\ell-1}$ représentent des nombres entiers différents de 1 à N, et aucun d'entre eux n'est équivalent à i ou j. Ensuite, Π_{ij}, la probabilité de la présence des unités i et j dans l'échantillon, s'obtient avec la formule suivante :

$$\Pi_{ij} = \sum_{k=1}^{n-2} \sum_{\ell=k+1}^{n-1} [\delta_{ij}(k,\ell) + \delta_{ji}(k,\ell)] + \sum_{k=1}^{n-1} [\delta_{ij}(k,n) + \delta_{ji}(k,n)]$$

$$= \sum_{k=1}^{n-1} \sum_{\ell=k+1}^{n} [\delta_{ij}(k,\ell) + \delta_{ji}(k,\ell)]$$

$$i = 1, 2, \ldots, N-1$$

$$j = i+1, \ldots, N. \qquad (3.1)$$

4 Échantillon rotatif

Supposons que nous voulions effectuer l'enquête dans les $m_\circ$ unités de premier degré (u.p.d.) d'une strate, pour certaines périodes précises. Il peut s'agir d'une période fixe prédéterminée ou du moment de l'épuisement d'au moins une u.p.d. Afin d'adopter ce genre de plan de

rotation, nous avons d'abord sélectionné $n = \sum_{t=0}^{T} m_t$ u.p.d., où m_o correspond à la définition ci-dessus, et m_t, au nombre d'u.p.d. nécessaires pour une rotation à la période t; t = 1, 2, 3, ... , T. À la période t = 0, nous prélevons un simple échantillon aléatoire de m_o dans n u.p.d. et, aux fins de rotation à la période de temps t, un simple échantillon aléatoire de m_t unités est sélectionné dans le reste de $n - (m_o + m_1 + \ldots + m_{t-1})$ des n unités sélectionnées initialement. Comme au départ, la probabilité de sélectionner une unité i correspondait à $\Pi_i = np_i$, et qu'à tout moment, la probabilité conditionnelle de sélectionner une unité (étant donné sa sélection initiale dans l'échantillonnage de premier degré) était équivalente à m_o/n, la probabilité inconditionnelle, Π'_i, de la présence de l'unité i dans l'échantillon final, correspond alors à

$$\begin{aligned} \Pi'_i &= \Pi_i \times \frac{m_o}{n} \\ &= np_i \times \frac{m_o}{n} \\ &= m_o p_i \qquad i = 1, 2, \ldots, N \end{aligned}$$

au besoin.

De même, la probabilité inconditionnelle, Π'_{ij}, de la présence des unités i et j dans l'échantillon s'obtient avec la formule suivante :

$$\Pi'_{ij} = \frac{m_o(m_o-1)}{n(n-1)} \Pi_{ij}$$

$$i = 1, 2, \ldots, N-1$$
$$j = i+1, i+2, \ldots, N.$$

où Π_{ij} s'obtient avec la formule (3.1).

Dans le plan de Fellegi [1], comme la probabilité de sélectionner une unité est proportionnelle à la taille de chacun des tirages successifs, nous sélectionnons donc des u.p.d. supplémentaires au moment de la rotation, pour l'échantillon de rotation.

5 Estimateur du total d'une population et de sa variance

La variable $s = \{i_1, i_2, \ldots, i_n\}$ désigne les n unités échantillonnées et la variable y_i désigne la valeur de la variable à l'étude pour l'unité i de la population; $i = 1, 2, \ldots, N$. Il faut estimer le total inconnu d'une population $Y = \sum_{i=1}^{N} y_i$ à partir des observations y_i pour $i \varepsilon s$. L'estimateur de Horvitz-Thompson [3] pour le total d'une population Y correspond à :

$$\hat{Y} = \frac{1}{n} \sum_{i \varepsilon s} \frac{y_i}{P_i} \tag{5.1}$$

et la variance de $\hat{Y}$ obtenue par Yates et Grundy [4], à

$$V(\hat{Y}) = \frac{1}{n^2} \sum_{i<j} \sum (\Pi_i \Pi_j - \Pi_{ij}) \left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^2 \tag{5.2}$$

où Π_i est la probabilité de la présence de l'unité i dans l'échantillon et Π_{ij} est la probabilité de la présence des unités i et j dans l'échantillon.

Un estimateur sans biais de $V(\hat{Y})$ est

$$\hat{V} = \frac{1}{n^2} \sum_{\substack{i<j \\ (i,j \varepsilon s)}} \sum \left(\frac{\Pi_i \Pi_j - \Pi_{ij}}{\Pi_{ij}}\right) \left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^2 . \tag{5.3}$$

La section suivante expose les résultats d'une étude empirique obtenus à partir de données tirées de Fellegi [1] et de Gray [2]. Les valeurs Π_{ij} tirées de la méthode de Fellegi [1] et de la méthode proposée pour des échantillons de taille 3 et 4 ont été présentées sous forme de tableau. Les valeurs Π_{ij} présentées sous forme de tableau, c'est-à-dire $\Pi_{ij} < \Pi_i \Pi_j$ pour toutes les paires (i, j) de la population, permettent de vérifier la non-négativité de l'estimateur de variance. Pour les deux méthodes, nous avons calculé les variances de $\hat{Y}$ et de $\hat{V}$ à partir d'échantillons de taille 3 et 4 issus de deux ensembles de données, c'est-à-dire celui de Fellegi [1] et celui de Gray [2].

6 Résultats empiriques

6.1 Exemple 1 : données tirées de Fellegi [1]

La population se compose de six unités d'échantillonnage primaires. Le tableau (6.1.1) présente les valeurs p_i et y_i.

Tableau (6.1.1)
Valeurs p_i et y_i de l'exemple (1)

i	1	2	3	4	5	6
p_i	0,10	0,14	0,17	0,18	0,19	0,22
y_i	0,60	0,98	1,53	2,16	2,85	4,18

$$Y = \sum_{i=1}^{N} y_i = 12{,}30$$

Le tableau (6.1.2) présente les « probabilités de travail » servant à la sélection d'un échantillon de taille 3 pour les deux plans. La colonne « nombre d'itérations » correspond au nombre d'itérations nécessaires pour atteindre la convergence* vers la solution. Il convient de mentionner que $p_i(k)$; i = 1, 2, ..., N représentent les « probabilités de travail » au k^e tirage; k = 1, 2, ..., n avec le plan de Fellegi [1], où $p_i(1) = p_i$; i = 1, 2, ..., N. Par ailleurs, q_i; i = 1, 2, ..., N représentent les « probabilités de travail » au énième tirage dans le plan proposé. Rappelons-nous que p_i; i = 1, 2, ..., N représentent les « probabilités de travail » à chacun des n - 1 premiers tirages du plan proposé.

Tableau (6.1.2)
« Probabilités de travail » pour la sélection de trois unités

Nombre d'itérations	i	1	2	3	4	5	6
-	p_i	0,100000	0,140000	0,170000	0,180000	0,190000	0,220000
-	$p_i(1)$	0,100000	0,140000	0,170000	0,180000	0,190000	0,220000
6	$p_i(2)$	0,090190	0,132410	0,167577	0,180157	0,193252	0,236414
9	$p_i(3)$	0,076367	0,119154	0,160684	0,177404	0,196167	0,270224
9	q_i	0,068868	0,113368	0,158820	0,177494	0,198613	0,282837

* La procédure d'itération se termine au moment où le changement de la valeur de chacun des éléments du vecteur de probabilité est inférieur ou égal à une donnée quantitative de 1.0E-8.

Nous remarquons, dans le tableau ci-dessus, que dans le plan de Fellegi [1], durant les tirages successifs, les « probabilités de travail » des unités 1, 2 et 3, qui correspondent aux trois plus petites unités de la population, diminuent, tandis que celles des unités 4, 5 et 6, qui correspondent aux plus grandes unités de la population, augmentent. Comme les « probabilités de travail » des n - 1 premiers tirages demeurent inchangées dans le plan proposé, celles des unités 1, 2 et 3 (les trois plus petites unités) au énième tirage (le troisième tirage dans ce cas-ci) sont plus petites que les « probabilités de travail » qui correspondent au plan de Fellegi [1]. Pour les unités 4, 5 et 6 (les trois plus grandes unités), les « probabilités de travail » sont plus grandes que celles qui correspondent au plan de Fellegi [1].

Le tableau suivant expose les valeurs de Π_{ij} dans les deux plans, pour un échantillon de taille 3. Les valeurs supérieures à la diagonale principale correspondent au plan de Fellegi [1] et celles inférieures à la principale diagonale, au plan proposé.

Tableau (6.1.3)
Valeurs de Π_{ij} pour un échantillon de taille 3

ij	1	2	3	4	5	6
1	X	0,086165	0,109898	0,118557	0,127675	0,157705
2	0,086163	X	0,161733	0,174315	0,187513	0,230274
3	0,109902	0,161742	X	0,221121	0,237550	0,289698
4	0,118557	0,174316	0,221111	X	0,255473	0,310534
5	0,127674	0,187510	0,237547	0,255479	X	0,331790
6	0,157704	0,230269	0,289699	0,310538	0,331791	X

Nous constatons, dans le tableau ci-dessus, que les valeurs de Π_{ij} des deux plans ne diffèrent pas jusqu'à quatre décimales et, puisque la variance est une fonction des valeurs de Π_{ij}, les deux plans seront donc aussi efficaces comme le montre le tableau suivant.

Tableau (6.1.4)
Variance de $\hat{Y}$ et de $\hat{V}$ pour les deux plans appliqués à un échantillon de taille 3

Plan de sélection	V($\hat{Y}$)	V($\hat{V}$)
Plan de Fellegi	3,8258	4,6166
Plan proposé	3,8259	4,6171

De la même manière, avec un échantillon de taille 4, les tableaux suivants permettent d'obtenir les « probabilités de travail », les valeurs de Π_{ij} et la variance de $\hat{Y}$ et de $\hat{V}$ pour les deux plans.

Tableau (6.1.5)
« Probabilités de travail » pour la sélection de quatre unités

Nombre d'itérations	i	1	2	3	4	5	6
-	p_i	0,100000	0,140000	0,170000	0,180000	0,190000	0,220000
-	$p_i(1)$	0,100000	0,140000	0,170000	0,180000	0,190000	0,220000
6	$p_i(2)$	0,090190	0,132410	0,167577	0,180157	0,193252	0,236414
9	$p_i(3)$	0,076367	0,119154	0,160684	0,177404	0,196167	0,270224
16	$p_i(4)$	0,051667	0,086649	0,130509	0,153692	0,184616	0,392867
17	q_i	0,033017	0,070222	0,121849	0,150012	0,187892	0,437008

Tableau (6.1.6)
Valeurs de Π_{ij} pour un échantillon de taille 4

ij	1	2	3	4	5	6
1	X	0,166245	0,216290	0,235994	0,256805	0,324666
2	0,167197	X	0,320554	0,348706	0,377519	0,466976
3	0,216261	0,320123	X	0,445971	0,478668	0,578517
4	0,235761	0,348406	0,446103	X	0,513248	0,616081
5	0,256432	0,377327	0,478869	0,513522	X	0,653760
6	0,324349	0,466948	0,578645	0,616208	0,653850	X

Tableau (6.1.7)
Variance de $\hat{Y}$ et de $\hat{V}$ pour les deux plans appliqués à un échantillon de taille 4

Plan de sélection	V($\hat{Y}$)	V($\hat{V}$)
Plan de Fellegi	1,5323	0,4672
Plan proposé	1,5269	0,4553

6.2 Exemple 2 : données tirées de Gray [2]

Dans cet exemple, la population est une strate de la Nouvelle-Écosse qui provient de l'Enquête sur la population active et qui comporte des caractéristiques fictives. La strate se compose de 10 unités d'échantillonnage primaires. Le tableau (6.2.1) présente les valeurs p_i et y_i.

Tableau (6.2.1)
Valeurs p_i et y_i de l'exemple (2)

i	1	2	3	4	5	6	7	8	9	10
p_i	0,0957	0,1043	0,1043	0,1006	0,0896	0,0881	0,0986	0,1055	0,1149	0,0984
y_i	10,06	10,35	10,38	9,57	9,30	8,96	10,00	10,50	11,33	9,55

$$Y = \sum_{i=1}^{N} y_i = 100,00$$

Comme dans l'exemple (1), nous avons calculé les « probabilités de travail », les valeurs de Π_{ij} et la variance de $\hat{Y}$ et de $\hat{V}$ pour les deux plans des échantillons de taille 3 et 4 tirés de données du tableau (6.2.1) ci-dessus. Les « probabilités de travail » ont affiché un comportement semblable au comportement vu dans l'exemple (1), et les valeurs de Π_{ij} des deux plans étaient identiques à 5 décimales dans les échantillons de taille 3 et 4. Des contraintes d'espace ont empêché de présenter les tableaux des « probabilités de travail » et des valeurs de Π_{ij}. Le tableau suivant présente le « nombre d'itérations » nécessaires pour obtenir les « probabilités de travail », de même que la variance de $\hat{Y}$ et de $\hat{V}$ pour des échantillons de taille 3 et 4.

Tableau (6.2.2)
« Nombre d'itérations » pour obtenir des « probabilités de travail », et variance de $\hat{Y}$ et de $\hat{V}$ pour les deux plans appliqués à des échantillons de taille 3 et 4

Plan de sélection	Taille de l'échantillon	Nombre d'itérations au tirage				V($\hat{Y}$)	V($\hat{V}$)
		1	2	3	4		
Plan de Fellegi	3	-	5	6		2,0509	2,7418
Plan proposé	3	-	-	6		2,0508	2,7418
Plan de Fellegi	4	-	5	6	7	1,3287	0,6647
Plan proposé	4	-	-	-	7	1,3287	0,6647

Nous constatons, pour les deux exemples de la présente étude, que les valeurs Π_{ij} des deux plans de sélection, c'est-à-dire le plan de Fellegi [1] et le plan proposé, sont presque identiques. Même si apparemment, les deux plans de sélection découlent d'un même plan de sondage, le choix entre les deux doit reposer sur la commodité opérationnelle. Le plan de Fellegi nécessite l'évaluation des « probabilités de travail » à chaque tirage sauf au premier, tandis que le plan proposé dans la présente étude

exige que nous les évaluions au dernier tirage seulement afin d'entraîner une réduction importante des calculs.

Remerciements

L'auteur tient à remercier monsieur G.B. Grey et monsieur M.P. Singh de la Division des méthodes d'enquête — recensement et enquêtes-ménages de Statistique Canada, de même que l'examinateur pour avoir fourni des commentaires constructifs.

Bibliographie

[1] Fellegi, I.P., "Sampling with Varying Probabilities Without Replacement: Rotating and Non-Rotating Samples", *Journal of the American Statistical Association*, Vol. 58 (1963), pp 183-201.

[2] Gray, G.B., "variance Components and Variance Function", *Proceedings of 'Statistics Canada 1971'*, pp 119-26.

[3] Horvitz, D.G. et Thompson, D.J., "A generalization of Sampling Without Replacement from a Finite Universe", *Journal of the American Statistical Association*, Vol. 47 (1952), pp 663-85.

[4] Yates, F. et Grundy, P.M., "selection Without Replacement from Within Strata with Probability Proportional to Size", *Journal of the Royal Statistical Society, Series B*, Vol. 15 (1953), pp 235-61.