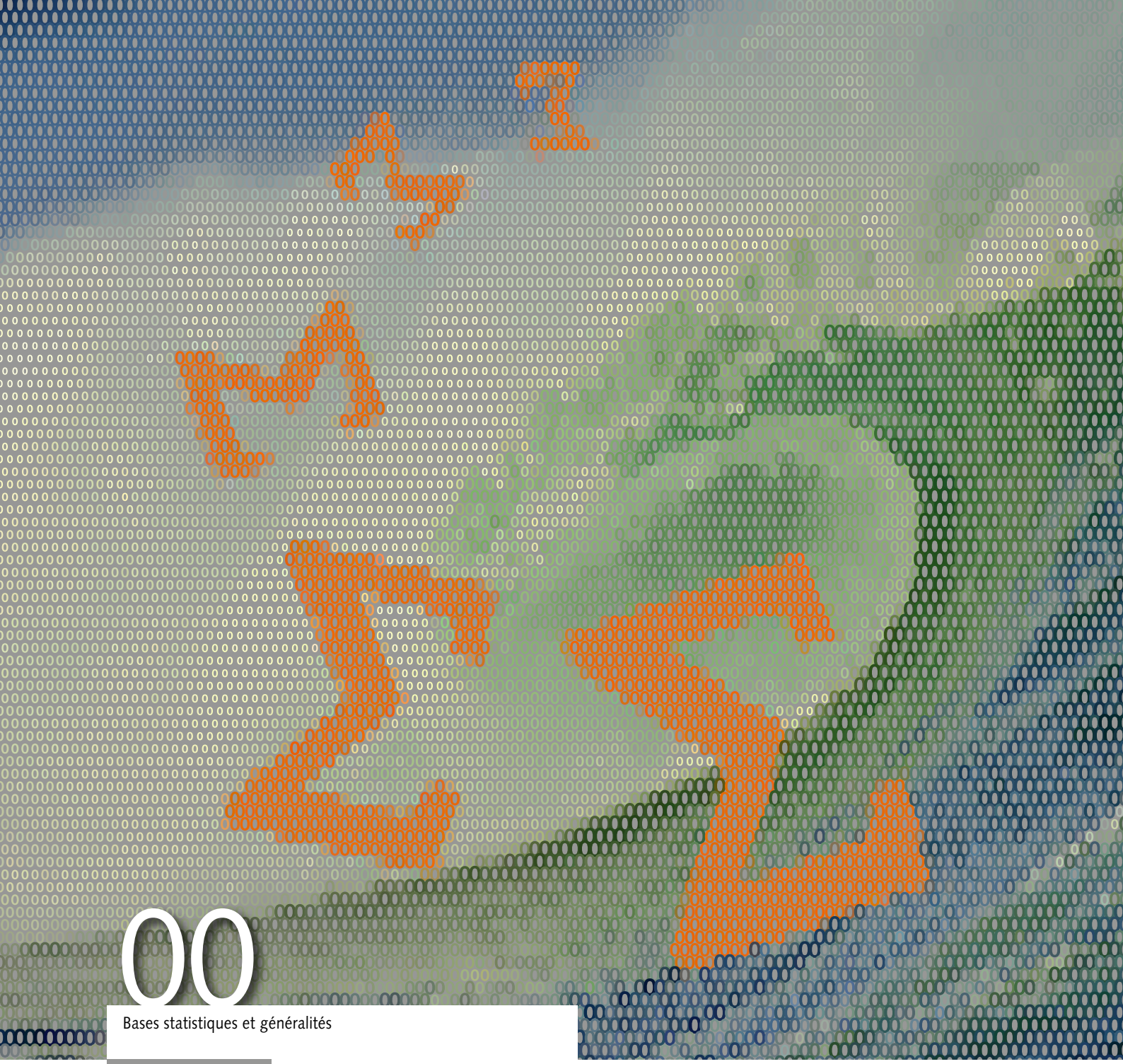


00

Bases statistiques et généralités

338-0073

Rapport de méthodes

Enquête sur la participation aux activités culturelles 2008

Plan de sondage et estimations

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
Office fédéral de la statistique OFS

Neuchâtel, 2014

La série «Statistique de la Suisse»
publiée par l'Office fédéral de la statistique (OFS)
couvre les domaines suivants:

0 Bases statistiques et généralités

1 Population

2 Espace et environnement

3 Vie active et rémunération du travail

4 Economie nationale

5 Prix

6 Industrie et services

7 Agriculture et sylviculture

8 Energie

9 Construction et logement

10 Tourisme

11 Mobilité et transports

12 Monnaie, banques, assurances

13 Protection sociale

14 Santé

15 Education et science

16 Culture, médias, société de l'information, sport

17 Politique

18 Administration et finances publiques

19 Criminalité et droit pénal

20 Situation économique et sociale de la population

21 Développement durable et disparités régionales et internationales

Rapport de méthodes

Enquête sur la participation aux activités culturelles 2008

Plan de sondage et estimations

Auteur Christian Panchard

Editeur Office fédéral de la statistique (OFS)

Office fédéral de la statistique (OFS)
Neuchâtel, 2014

Editeur: Office fédéral de la statistique (OFS)
Complément d'information: Christian Panchard, tél. 058 463 61 81, e-mail: christian.panchard@bfs.admin.ch
Réalisation: Section Méthodes statistiques, OFS
Diffusion: Office fédéral de la statistique, CH-2010 Neuchâtel
tél. 058 463 60 60 / fax 058 463 60 61 / e-mail: order@bfs.admin.ch
Internet: www.statistique.admin.ch
Numéro de commande: 338-0073
Prix: gratuit
Série: Statistique de la Suisse
Domaine: 0 Bases statistiques et généralités
Langue du texte original: Français
Page de couverture: OFS; concept: Netthoevel & Gaberthüel, Bienne; photo: © NorthShoreSurfPhotos – Fotolia.com
Graphisme/Layout: Section DIAM, Prepress/Print
Copyright: OFS, Neuchâtel 2014
La reproduction est autorisée, sauf à des fins commerciales,
si la source est mentionnée
ISBN: 978-3-303-00517-0

Table des matières

Préambule 5

Résumé 5

1 Introduction 7

2 Plan de sondage 8

2.1 Population "cible" et cadre de sondage 8

2.2 Taille de l'échantillon et stratification 8

2.3 Allocation et plan de sondage 8

3 Tirage et activation de l'échantillon 10

4 Pondération et estimation 11

4.1 Probabilité d'inclusion des ménages et poids de sondage "ménage" 11

4.2 Correction de la non-réponse des ménages 12

4.3 Le poids "personne" 14

4.4 Calage sur marges 15

5 Estimation de la variance 21

6 Résultats 23

Références 27

Rapports de méthodes de la section méthodes statistiques de l'OFS 29

Préambule

À la demande de la section POKU (Politique, culture et médias) de l'Office fédéral de la statistique, la section METH (Méthodes statistiques) lui a fourni un soutient pour l'élaboration d'un plan de sondage, la pondération des données des répondants et les estimations de l'enquête sur la participation aux activités culturelles. On trouvera dans ce rapport une description de ces étapes successives. Les travaux ont été réalisés par Christian Panchard et Paul-André Salamin.

Un grand merci à Stéphanie Vanhooydonck et Olivier Moeschler de la section POKU pour leur aide et leurs remarques tout au long de ce travail. Merci également à Daniel Assoulin de la section METH pour la relecture de ce rapport et ses remarques toujours pertinentes.

Résumé

Ce rapport documente le plan de sondage, l'allocation et le tirage de l'échantillon, ainsi que les méthodes d'estimation utilisées pour la statistique concernant la participation aux activités culturelles. Cette statistique se base sur une enquête auprès des ménages résidents en Suisse. Le plan d'échantillonage est conçu de manière à pouvoir mettre en évidence les différences culturelles relatives aux régions linguistiques de la Suisse. Ces dernières constituent donc les strates du tirage et l'allocation choisie tient compte de la précision et de la non-réponse au niveau de ces strates. L'échantillon se compose de $7'423$ personnes pour lesquelles on obtient $4'346$ répondants. La pondération de ces répondants comporte une correction de la non-réponse et un calage, de telle sorte qu'elle permet des extrapolations de l'échantillon à la population de la Suisse. Ces extrapolations sont accompagnées d'une estimation de variance.

1 Introduction

L'enquête suisse sur la participation aux activités culturelles est un projet mené en commun par l'Office fédéral de la Statistique (OFS) et l'Office fédéral de la Culture (OFC). Sa base légale se trouve dans le projet de loi sur l'encouragement de la culture du Conseil fédéral à l'attention du Parlement. Ce projet ayant abouti en 2009 à la loi sur l'encouragement de la culture (LEC), l'assemblée fédérale y arrête, entre autre, sous l'article 30 (Statistique et évaluation), que "L'Office fédéral de la statistique tient une statistique culturelle".

Le but de l'enquête est de développer une statistique sur les pratiques culturelles et de compléter les autres statistiques produites actuellement à l'OFS, et ce en collaboration avec l'OFC qui se charge du financement de l'enquête. Cette nouvelle statistique doit permettre de décrire les pratiques culturelles telles que la fréquentation d'institution culturelles (concerts, théatres, cinémas, visites du patrimoine, etc.), la pratique d'activités en amateur (chanter, jouer d'un instrument, peindre, etc.) ou encore l'utilisation des médias (TV, radio, presse, livres). Elle se base sur une enquête téléphonique menée fin 2008 auprès d'un échantillon net de $4'346$ personnes de 15 ans et plus de la population en Suisse.

Pour plus de détails voir [5] et [10]

2 Plan de sondage

2.1 Population "cible" et cadre de sondage

La population "cible" de l'enquête sur la participation aux activités culturelles était constituée des personnes de la population résidente permanente de Suisse, âgées de 15 ans et plus. En 2008, année du tirage de l'échantillon, l'OFS ne disposait, pour les enquêtes auprès des ménages, que de la base CASTEM (CAdre de Sondage pour le Tirage d'Échantillon de Ménage), construite sur la base des livraisons par tous les opérateurs téléphoniques de Suisse de tous leurs numéros de téléphones fixes. Il s'agissait donc d'une base de numéros de téléphone qu'on identifie, après sélection des numéros de ménages privés, avec l'ensemble des ménages de Suisse. La nature de cette base et les procédures de sa construction la rende sujette à de la sous-couverture (tous les ménages n'ont pas de numéros fixes) et à de la sur-couverture (l'identification des numéros de ménages privés est parfois difficile à effectuer et il est possible que certains numéros d'entreprises ou de ménages collectifs se retrouvent dans la base). Ces deux cas de défaut de couverture sont corrigés lors de la pondération des répondants à l'enquête. La taille de la base lors du tirage de l'échantillon en juillet 2008 était de $3'333'274$ ménages.

L'enquête portant sur des personnes et non des ménages, le plan de sondage devait être un plan à deux degrés, ce qui signifie qu'il fallait, après avoir tiré aléatoirement un échantillon de ménage, choisir encore aléatoirement une personne "cible" parmi toutes celles du ménage qui ont plus de 15 ans.

2.2 Taille de l'échantillon et stratification

La taille de l'échantillon net était fixée à environ $4'500$ personnes répondantes pour toute la Suisse, ce qui permet de déterminer une taille d'échantillon brut d'environ $6'820$ ménages tirés dans CASTEM, sachant que le taux de ménages valides (numéros de téléphones de vrais ménages et non pas d'entreprise, de home, etc.) est, pour ce type d'échantillon, d'environ 88% de la totalité des ménages et le taux d'interviews entièrement complétées d'environ 75% des ménages valides (taux moyens obtenus lors d'enquêtes précédentes).

La stratification retenue pour cette enquête était composée des régions linguistiques, ce qui correspond exactement aux domaines d'analyse envisagés et ainsi tenait bien compte de possibles différences culturelles. Les tailles de la base CASTEM pour les strates, en juillet 2008, étaient de $2'371'352$ ménages pour la région germanophone (71%), de $798'001$ pour la partie francophone (24%) et de $163'921$ pour la partie italophone (5%). Par la suite la strate correspondant à la région germanophone sera notée "$\mathcal{D}$", celle correspondant à la strate francophone "$\mathcal{F}$" et celle correspondant à la strate italophone "$\mathcal{I}$".

2.3 Allocation et plan de sondage

Pour l'allocation de la taille de l'échantillon brut aux différentes strates, deux modèles ont été étudiés. Le premier proposait une allocation proportionnelle aux tailles des strates de CASTEM ($4'851$ ménages pour la strate $\mathcal{D}$, $1'632$ ménages pour la strate $\mathcal{F}$ et 335 ménages pour la

strate $\mathcal{I}$), alors que le second introduisait une surreprésentation de la strate $\mathcal{I}$ qui la rendait approximativement de même taille que la strate $\mathcal{F}$ ($1'364$ ménages), tout en répartissant le reste de l'échantillon brut proportionnellement aux strates $\mathcal{D}$ ($4'081$ ménages) et $\mathcal{F}$ ($1'373$ ménages).

Le choix entre ces deux modèles a été déterminé par une étude de la précision d'une estimation Horvitz-Thompson d'une proportion p, pour une taille d'échantillon net m à laquelle on peut s'attendre avec la taille d'échantillon brut fixée n, au vu d'enquêtes similaires menées par l'OFS (par exemple les enquêtes sur la santé SSE). Cette étude considérait le coefficient de variation de la proportion $p = \frac{1}{2}$, dans le cas d'un tirage aléatoire simple. Ce coefficient est alors donné par :

$$cv = \frac{\sqrt{\frac{p\cdot(1-p)}{m}}}{p} = \frac{\sqrt{\frac{\frac{1}{2}\cdot(1-\frac{1}{2})}{m}}}{\frac{1}{2}} = \frac{1}{\sqrt{m}}.$$

Les résultats de cette étude se retrouvent dans les Tableaux 1, page 9, et 2, page 9.

TABLEAU 1 - Allocation proportionnelle.

Strate	n	m	cv (en %)
$\mathcal{D}$	4'851	3'202	1.8
$\mathcal{F}$	1'632	1'077	3.0
$\mathcal{I}$	335	221	6.7

TABLEAU 2 - Surreprésentation de la strate I.

Strate	n	m	cv (en %)
$\mathcal{D}$	4'081	2'694	1.9
$\mathcal{F}$	1'373	906	3.3
$\mathcal{I}$	1'364	900	3.3

Avec une surreprésentation de la strate $\mathcal{I}$ on obtient une estimation bien plus précise pour cette strate (3.3% contre 6.7%) sans pour autant beaucoup détériorer les précisions des autres strates. Cet argument a été décisif pour le choix du deuxième modèle.

3 Tirage et activation de l'échantillon

Le tirage simple et stratifié de l'échantillon a été effectué en août 2008 dans la deuxième base CASTEM de l'année en cours. Aux $6'818$ ménages de l'échantillon brut on a ajouté une réserve de $3'782$ ménages qui étaient répartis en 16 sous-réserves, afin de permettre d'éventuelles activations successives selon l'évolution des taux de réponse lors du déroulement de l'enquête. Ainsi, des $10'000$ ménages tirés, $7'423$ ont été activés au final ($4'442$ pour strate $\mathcal{D}$, $1'496$ pour strate $\mathcal{F}$ et $1'485$ pour strate $\mathcal{I}$).

Le Tableau 3, page 10, donne l'évolution des tailles de l'échantillon à chaque étape de l'enquête, à savoir, en partant de l'échantillon brut des ménages activés, le nombre de ménages valides (i.e. appartenant bien à la population "cible" des ménages privés), puis le nombre de répondants "ménage" (i.e. le nombre de ménages qui acceptent de participer à l'enquête), et enfin, après le choix d'une personne "cible" de plus de 15 ans dans ces ménages, le nombre de répondants "personne" (i.e. le nombre de ménages pour lesquels la personne "cible" accepte également de participer, ce qui correspond aussi au nombre d'interviews entièrement complétées).

TABLEAU 3 - Évolution de l'échantillon (les % se réfèrent à la ligne précédente).

	$\mathcal{D}$	$\mathcal{F}$	$\mathcal{I}$	Total
ménages activés	4'442	1'496	1'485	7'423
ménages valides	4'049 (91%)	1'318 (88%)	1'250 (84%)	6'617 (89%)
répondants "ménage"	3'098 (77%)	974 (74%)	961 (77%)	5'033 (76%)
répondants "personne"	2'698 (87%)	848 (87%)	800 (83%)	4'346 (86%)

Le taux de ménage valides était de 89% (91% pour la strate $\mathcal{D}$, 88% pour la strate $\mathcal{F}$ et 84% pour la la strate $\mathcal{I}$), ce qui correspond assez bien aux 88% estimés lors de l'élaboration du plan d'échantillon. Par contre le taux d'interviews entièrement complétées n'était que de 66% des ménages valides (67% pour la strate $\mathcal{D}$, 64% pour la strate $\mathcal{F}$ et 64% pour la strate $\mathcal{I}$), taux nettement inférieur aux 75% estimés, ce qui explique qu'on ait dû activer une partie des réserves.

Les 4'346 répondants "personne" constituent l'échantillon net de l'enquête qui sera utilisé pour les estimations, après pondération.

4 Pondération et estimation

4.1 Probabilité d'inclusion des ménages et poids de sondage "ménage"

Soient U la population de la base CASTEM, S l'échantillon brut des ménages activés, $\widetilde{S}$ l'échantillon des ménages valides de S et R l'échantillon net des répondants "ménage".

Soient encore U_h les populations des strates, avec $h \in \mathcal{D}, \mathcal{F}, \mathcal{I}$, et les échantillons correspondants S_h, $\widetilde{S}_h$ et R_h.

On note N la taille de U (N_h pour U_h), n et $\tilde{n}$ celles de S et $\widetilde{S}$ (n_h pour S_h, et $\tilde{n}_h$ pour $\widetilde{S}_h$), et m celle de R (m_h pour R_h).

Le Tableau 4 (page 11) récapitule les valeurs de ces différentes tailles déjà données dans les paragraphes précédents :

TABLEAU 4 - Les tailles des échantillons.

h	N_h	n_h	$\tilde{n}_h$	m_h
$\mathcal{D}$	2'371'352	4'442	4'049	3'098
$\mathcal{F}$	798'001	1'496	1'318	974
$\mathcal{I}$	163'921	1'485	1'250	961
Total	3'333'274	7'423	6'617	5'033

La première étape de la pondération doit tenir compte du plan de sondage (voir [9], page 171 et suivantes). La probabilité d'inclusion d'un ménage $k \in U_h$

$$\pi_{Ik} = \frac{n_h}{N_h}\ , \quad h \in \mathcal{D}, \mathcal{F}, \mathcal{I}$$

permet d'exprimer le poids de sondage de ce ménage :

$$w_{Ik} = \frac{1}{\pi_{Ik}} = \frac{N_h}{n_h}.$$

On obtient ainsi un unique poids de sondage "ménage" w_{Ik} par strate h, $h \in \mathcal{D}, \mathcal{F}, \mathcal{I}$ (Tableau 5, page 12).

TABLEAU 5 - Les poids "ménage".

h	w_{Ik}
$\mathcal{D}$	533.8
$\mathcal{F}$	533.4
$\mathcal{I}$	110.4

4.2 Correction de la non-réponse des ménages

La seconde étape de la pondération consiste à corriger les effets de la non-réponse des ménages. Cette correction s'effectue au niveau des strates. En effet, CASTEM contenant comme information le nom, le prénom, l'adresse et le numéro de téléphone , seule la localisation du ménage peut être utilisée pour une telle correction. Elle peut être réalisée au niveau des cantons, des grandes régions ou des régions linguistiques (i.e. les strates). Le nombre de répondants "ménage" des cantons étant souvent petit (par exemple 14 pour Glaris ou 27 pour le Jura) et les grandes régions ne recouvrant pas les régions linguistiques, le choix s'est finalement reportés sur les régions linguistiques.

Deux options s'offrent pour une telle correction. Il est possible de considérer que la non-réponse correspond soit au nombre de non-répondants $S \setminus R$ par rapport à la totalité de l'échantillon brut activé S, soit au nombre de non-répondants $\widetilde{S} \setminus R$ par rapport à la totalité de l'échantillon des ménages valides $\widetilde{S}$. La pemière option suppose qu'à la non-réponse au sens strict (basée uniquement sur les ménages valides qui seuls participent à l'enquête) se mêle une partie de la correction de sur-couverture, S contenant les ménages invalides, alors que la deuxième ne tient compte que de la non-réponse au sens strict.

4.2.1 Correction par rapport à l'échantillon brut S.

Dans le cas de la première variante de correction de la non-réponse, si l'on note

$$\check{\rho}_h = \frac{m_h}{n_h}$$

la probabilité de répondre d'un ménage $k \in S_h$, $h \in \mathcal{D}, \mathcal{F}, \mathcal{I}$, on peut écrire la correction pour la non-réponse :

$$\check{\kappa}_h = \frac{1}{\check{\rho}_h} = \frac{n_h}{m_h},$$

ce qui permet d'exprimer la probabilité d'inclusion d'un ménage $k \in S_h$ dans l'échantillon net des répondants "ménage" $R \subset S_h \subset U_h$:

$$\check{\pi}_{Ik} = \pi_{Ik} \cdot \check{\rho}_h = \frac{n_h}{N_h} \cdot \frac{m_h}{n_h} = \frac{m_h}{N_h},$$

et le poids "ménage" corrigé pour la non-réponse :

$$\breve{w}_{Ik} = w_{Ik} \cdot \breve{\kappa}_h = \frac{N_h}{n_h} \cdot \frac{n_h}{m_h} = \frac{N_h}{m_h}.$$

4.2.2 Correction par rapport à l'échantillon des valides $\widetilde{S}$.

Pour la deuxième variante, de manière identique, en notant

$$\tilde{\rho}_h = \frac{m_h}{\tilde{n}_h}$$

la probabilité de répondre d'un ménage $k \in \widetilde{S}_h$, $h \in \mathcal{D}, \mathcal{F}, \mathcal{I}$, on peut écrire la correction pour la non-réponse :

$$\tilde{\kappa}_h = \frac{1}{\tilde{\rho}_h} = \frac{\tilde{n}_h}{m_h},$$

et la probabilité d'inclusion d'un ménage $k \in \widetilde{S}_h$ dans l'échantillon des ménages valides $R \subset \widetilde{S}_h \subset U_h$:

$$\tilde{\pi}_{Ik} = \pi_{Ik} \cdot \tilde{\rho}_h = \frac{n_h}{N_h} \cdot \frac{m_h}{\tilde{n}_h}.$$

On obtient alors le poids "ménage" corrigé pour la non-réponse :

$$\tilde{w}_{Ik} = w_{Ik} \cdot \tilde{\kappa}_h = \frac{N_h}{n_h} \cdot \frac{\tilde{n}_h}{m_h}\ , \quad h \in \mathcal{D}, \mathcal{F}, \mathcal{I}.$$

4.2.3 Choix de la correction.

Les poids "ménage" corrigés pour la non-réponse $\breve{w}_h$ et $\tilde{w}_h$ sont donnés dans le tableau Tableau 6, page 13.

TABLEAU 6 - Les poids "ménage" $\breve{w}_{Ik}$, corrigés par rapport à S, et $\tilde{w}_{Ik}$, corrigés par rapport à $\widetilde{S}$.

h	$\breve{w}_{Ik}$	$\tilde{w}_{Ik}$
$\mathcal{D}$	878.9	801.2
$\mathcal{F}$	941.0	829.1
$\mathcal{I}$	204.9	172.5

Alors que, pour chacun des ces poids, les différences relatives entre les strates sont approximativement les mêmes, les poids $\check{w}_{Ik}$ sont plus élevés que les poids $\tilde{w}_{Ik}$, ce qui se traduit aussi par le fait que la somme des poids $\check{w}_{Ik}$ sur R donne $3'333'274$ et celle des poids $\tilde{w}_{Ik}$ $3'002'582$. La première de ces sommes, qui toutes deux donnent une estimation du nombre de ménage de la population "cible", se rapproche plus des $3'248'200$ ménages en Suisse en 2005. Toutefois, comme les poids $\tilde{w}_{Ik}$ correspondent à une correction de la non-réponse au sens strict, on les préfère aux poids $\check{w}_{Ik}$. La sous-estimation du nombre de ménages, qui correspond à la sous-couverture évoquée au paragraphe 2.1, sera corrigée dans la cadre du calage qui constitue une des étapes suivantes de la pondération.

4.3 Le poids "personne"

Soit t_k le nombre de personnes de plus de 15 ans du ménage k. La probabilité que la personne "cible" i de ce ménage soit sélectionnée s'écrit $\pi_{i|k} = \frac{1}{t_k}$ et on peut alors exprimer la probabilité de i de se retrouver dans l'échantillon des répondants "personne" :

$$\pi_i = \tilde{\pi}_{Ik} \cdot \pi_{i|k} = \frac{n_h}{N_h} \cdot \frac{m_h}{\tilde{n}_h} \cdot \frac{1}{t_k} \ , \quad h \in \mathcal{D}, \mathcal{F}, \mathcal{I},$$

et son poids "personne" :

$$w_i = \frac{1}{\pi_i} = \frac{N_h}{n_h} \cdot \frac{\tilde{n}_h}{m_h} \cdot t_k = \tilde{w}_{Ik} \cdot t_k \ ,$$

poids ne tenant pas compte de la non-réponse qui sera corrigée lors du calage de l'étape suivante.

La répartition de ces poids "personne" w_i, pour i dans l'échantillon des répondants "personne" $\widetilde{R}$, de taille $\widetilde{m}$, est donnée par le tableau Tableau 7, page 14, et la Figure 1, page 15.

TABLEAU 7 - Répartition des poids "personne" w_i.

$\widetilde{m}$	*Moyenne*	*CV*	*Total*
4346	1281.66	60.0112	5570085.71

Quantiles.

P0	*P1*	*P5*	*P10*	*P25*	*P50*	*P75*	*P90*	*P95*	*P99*	*P100*
172.5	172.5	172.5	345.0	801.2	1602.3	1602.3	2403.5	2487.2	3316.3	5608.2

On peut constater que la somme totale des poids "personne" w_i est égale à $5'570'086$, ce qui est assez éloigné de la population résidentes permanentes de Suisse, qui était de $6'183'275$ habitants selon la statistique de l'état annuel de la population 2007 (ESPOP 2007). Cette différence

provient en grande partie de la sous-couverture de la base CASTEM, ainsi que de l'absence de correction de la non-réponse des personnes "cible". Ces deux causes de sous-estimation sont corrigées dans l'étape suivante de la pondération.

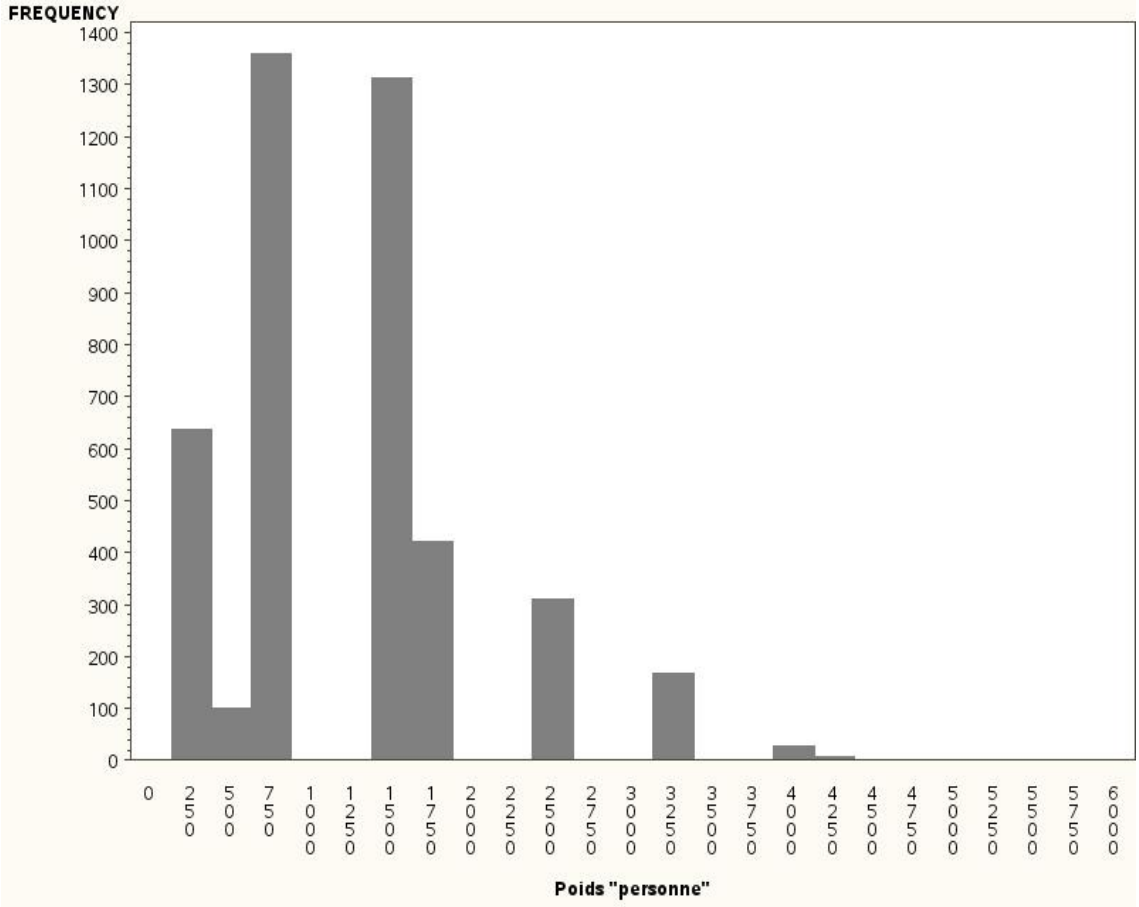


FIGURE 1 Répartition des poids de sondage personne.

4.4 Calage sur marges

Soient P l'ensemble de la population de Suisse selon ESPOP 2007. Le but de la pondération est de permettre l'estimation du total d'une variable y_i, $i \in P$:

$$Y = \sum_{i \in P} y_i.$$

Après le calcul d'un poids "personne" w_i, cette estimation peut être obtenue par l'estimateur de Horvitz-Thompson :

$$\hat{Y}_{HT} = \sum_{i \in \widetilde{R}} w_i \cdot y_i.$$

Lorsque y_i est la fonction indicatrice d'une sous-population de P ($y_i = 1$ sur la sous-population, $= 0$ ailleurs), l'estimation $\sum_{i \in \widetilde{R}} w_i \cdot y_i$ devrait donner la taille de la sous-population considérée. Avec les poids w_i ce n'est généralement pas le cas, comme on l'a vu, par exemple, au paragraphe 4.3 pour l'estimation de la taille de P par la somme totale des poids "personne" w_i ($\sum_{i \in P} w_i \cdot y_i$, avec $y_i = 1, \forall i \in P$). L'objectif de l'étape de calage sur marges est de trouver des nouveaux poids $\bar{w}_i$, aussi proche que possible de w_i, qui permettent de retrouver les tailles de certaines sous-populations choisies de P.

Soit x_i un vecteur de p variables auxilaires tel que $\sum_{i \in P} x_i$ soit connu et soit une fonction de distance G d'argument $z = \frac{\bar{w}_i}{w_i}$, mesurant la distance entre $\bar{w}_i$ et w_i, positive convexe et telle que $G(1) = G'(1) = 0$.

La méthode de calage sur marges consiste à résoudre un problème d'optimisation, à savoir chercher le minimum

$$\min_{\bar{w}_i, i \in \tilde{R}} \sum_{i \in \widetilde{R}} w_i \cdot G(\frac{\bar{w}_i}{w_i})$$

sous les containtes

$$\sum_{i \in \widetilde{R}} \bar{w}_i \cdot x_i = \sum_{i \in P} x_i = x_P.$$

Ce problème est résolu en introduisant un vecteur de multiplicateurs de Lagrange $\lambda' = (\lambda_1 \cdots \lambda_p)$ qui permet d'écrire le Lagrangien :

$$L = \sum_{i \in \tilde{R}} w_i \cdot G(\frac{\bar{w}_i}{w_i}) - \lambda' \cdot (\sum_{i \in \tilde{R}} \bar{w}_i \cdot x_i - x_P).$$

Les conditions du premier ordre donnent :

$$\bar{w}_i = w_i \cdot F(x_i \lambda_i)$$

où F est la fonction réciproque de la dérivée de la fonction G.

Le vecteur λ' est déterminé par la résolution du système non linéaire :

$$\sum_{i \in \widetilde{R}} w_i \cdot F(x_i \lambda_i) \cdot x_i = x_P.$$

Une résolution numérique de ce sytème est obtenue par la méthode itérative de Newton, en utilisant la macro SAS CALMAR développée par l'INSEE.

Comme fonction de distance on choisit $G(z) = z \cdot \log z - z + 1$, ce qui correspond à la méthode "raking ratio", méthode qui donne des poids $\bar{w}_i$ toujours positifs.

Le vecteur x_i est construit à l'aide des indicatrices des sous-populations déterminées par le sexe des répondants, leur classe d'âge, leur état civil (réduit à *marié(e) / célibataire*) et leur nationalité (réduite à *Suisse / étranger*). Le choix de ces sous-populations répond à des besoins concernant les études envisagées et leurs totaux de référence sont donnés par ESPOP 2007 (2008 n'étant pas encore disponible lors du calage). Ces derniers se trouvent dans le Tableau 8, page 18.

Pour plus de détails voir [3] et [8].

La Figure 2, page 19, montre la tendance du calage à augmenter les poids. La répartition des poids corrigés $\bar{w}_i$ est donnée par le tablau Tableau 9, page 19, ainsi que par la Figure 3 (page 20).

TABLEAU 8 - Les totaux du calage (les % sont donnés par rapport au total).

	Échantillon				Totaux ESPOP	
	$\widetilde{m}$	%	$\sum_{\widetilde{R}} w_i$	%	N	%
Total	4'346		5'570'086		6'195'328	
Sous-population						
Sexe						
Homme	1'888	43.4%	2'569'203	46.1%	3'032'059	48.9%
Femme	2'458	56.6%	3'000'883	53.9%	3'163'269	51.1%
Classe d'âge						
15 - 19 *ans*	194	4.5%	445'222	8.0%	435'480	7.0%
20 - 24 *ans*	180	4.1%	324'947	5.8%	426'574	6.9%
25 - 34 *ans*	568	13.1%	727'380	13.1%	975'538	15.7%
35 - 44 *ans*	893	20.5%	1'097'771	19.7%	1'219'358	19.7%
45 - 54 *ans*	720	16.6%	996'786	17.9%	1'093'828	17.7%
55 - 64 *ans*	713	16.4%	889'290	16.0%	905'657	14.6%
65 *ans et plus*	1'078	24.8%	1'088'690	19.5%	1'138'893	18.4%
État civil						
marié(e)	2'212	50.9%	2'462'958	44.2%	2'782'153	44.9%
célibataire	2'134	49.1%	3'107'128	55.8%	3'413'175	55.1%
Nationalité						
Suisse	3'652	84.0%	4'715'478	84.7%	4'887'234	78.9%
Étranger	694	16.0%	854'608	15.3%	1'308'094	21.1%

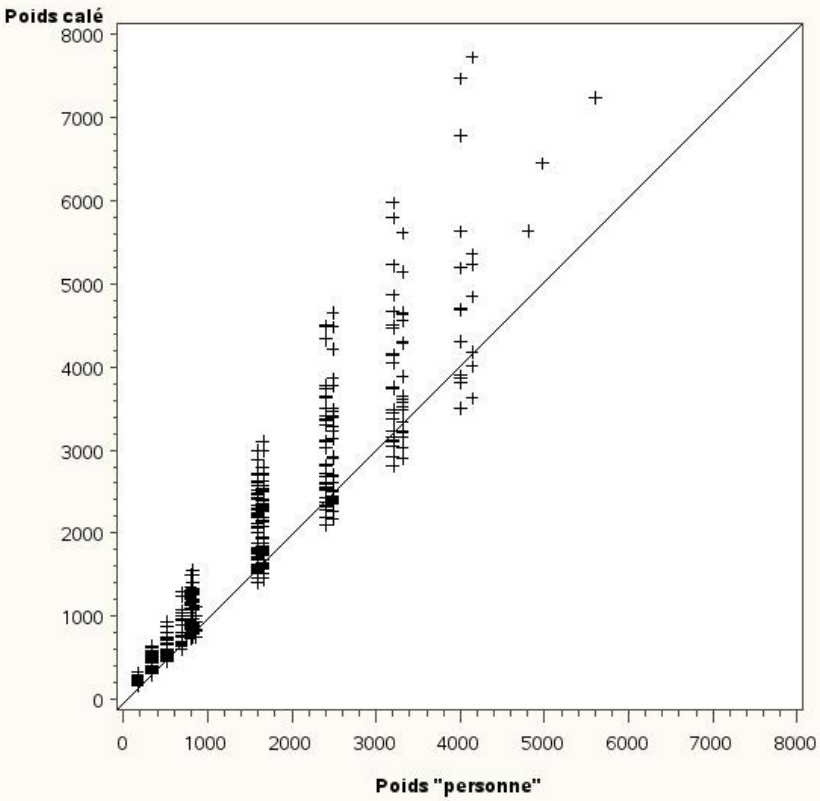


FIGURE 2 Effet du calage sur les poids.

TABLEAU 9. Répartition des poids $\bar{w}_i$.

n	*Moyenne*	*CV*	*Total*
4346	1425.5	63.8	6195328

Quantiles.

P0	*P1*	*P5*	*P10*	*P25*	*P50*	*P75*	*P90*	*P95*	*P99*	*P100*
157.5	169.8	235.3	362.7	797.4	1463.3	1762.7	2544.5	3106.1	4213.2	7733.3

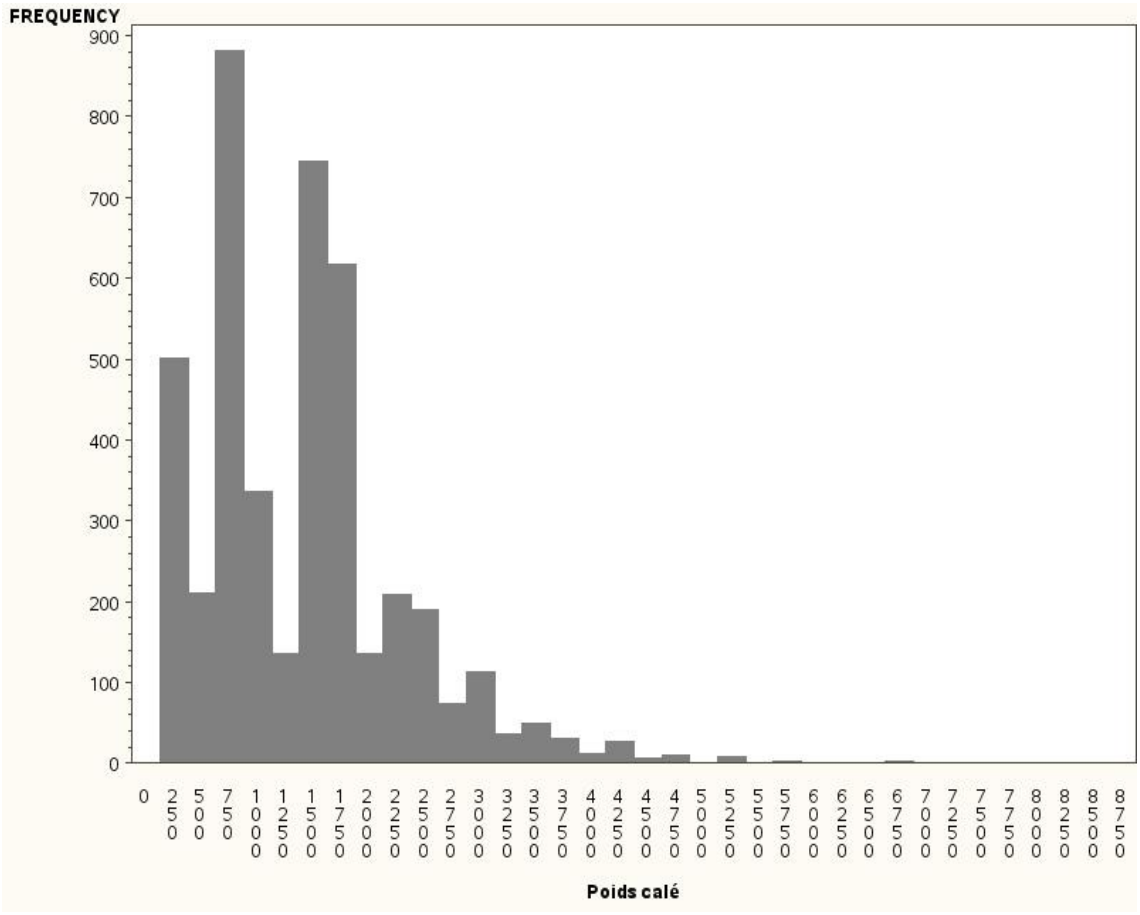


FIGURE 3 Répartition des poids calés.

5 Estimation de la variance

Cette section suit en grande partie le rapport de méthodes de Andrade et Salamin [1].

Rappelons que l'on cherche à estimer le total d'une variable y_i, $i \in P$:

$$Y = \sum_{i \in P} y_i.$$

L'estimateur de régression généralisé d'un tel total est donné par :

$$\widehat{Y} = \widehat{Y}_{HT} + (x_P - \hat{x}_{HT})^t \, \widehat{\beta}_{HT},$$

avec

$$\widehat{Y}_{HT} = \sum_{i \in \widetilde{R}} w_i \cdot y_i, \;\; x_P = \sum_{i \in P} x_i = \sum_{i \in \widetilde{R}} \bar{w}_i \cdot x_i, \;\; \hat{x}_{HT} = \sum_{i \in \widetilde{R}} w_i \cdot x_i$$

et

$$\widehat{\beta}_{HT} = \Big(\sum_{i \in \widetilde{R}} w_i \cdot x_i \cdot x_i^t \Big)^{-1} (\sum_{i \in \widetilde{R}} w_i \cdot x_i \cdot y_i).$$

Le développement en série de Taylor de premier ordre de $\hat{Y}$ nous donne :

$$\widehat{Y} \approx \widehat{Y}_{HT} + (x_P - \hat{x}_{HT})^t \cdot \beta = \widehat{Y}_{approx},$$

avec β obtenu par la méthode des moindres carrés :

$$\beta = \Big(\sum_{i \in P} x_i \cdot x_i^t \Big)^{-1} (\sum_{i \in P} x_i \cdot y_i),$$

Cette approximation de $\widehat{Y}$ par le développement de Taylor peut aussi s'écrire :

$$\widehat{Y}_{approx} = x_P^t \cdot \beta + \sum_{i \in \widetilde{R}} w_i \cdot \Big(y_i - x_i^t \cdot \beta \Big) = \sum_{i \in P} x_i^t \cdot \beta + \sum_{i \in \widetilde{R}} w_i \cdot e_i,$$

où

$$e_i = y_i - x_i^t \cdot \beta$$

sont les résidus sur la population.

Sa variance est donnée par :

$$\begin{aligned}\mathrm{Var}\Big(\widehat{Y}_{approx}\Big) &= \mathrm{Var}\Big(\sum_{i\in\widetilde{R}} w_i\cdot e_i\Big) = \sum_{i,i'\in P}\Delta_{ii'}\cdot w_i\cdot w_{i'}\cdot e_i\cdot e_{i'}\\ &= \sum_{i,i'\in P}\Delta_{ii'}\cdot w_i\cdot\Big(y_i - x_i^t\beta\Big)\cdot w_{i'}\cdot\Big(y_{i'} - x_{i'}^t\beta\Big),\end{aligned}$$

avec $\Delta_{ii'} = \pi_{ii'} - \pi_i\cdot\pi_{i'}$ où $\pi_i = \frac{1}{w_i}$ et $\pi_{ii'}$ désigne la probabilité $P\Big(i, i'\in\widetilde{R}\Big)$.

Un estimateur de cette variance peut être calculé par :

$$\begin{aligned}\widehat{\mathrm{Var}\Big(\widehat{Y}\Big)} &= \sum_{i,i'\in\tilde{R}}\frac{\Delta_{ii'}}{\pi_{ii'}}\cdot w_i\cdot\Big(y_i - x_i^t\hat{\beta}_{HT}\Big)\cdot w_{i'}\cdot\Big(y_{i'} - x_{i'}^t\hat{\beta}_{HT}\Big)\\ &= \sum_{i,i'\in\tilde{R}}\frac{\Delta_{ii'}}{\pi_{ii'}}\cdot w_i\cdot w_{i'}\cdot\hat{e}_i\cdot\hat{e}_{i'}.\end{aligned}$$

6 Résultats

Soient Q une variable catégorielle associée à une question de l'enquête, et les variables indicatrices δ_j associées aux modalités j de la variable Q :

$$\forall i \in \widetilde{R},\ \delta_{ij} = \begin{cases} 1, & \text{si } Q_i = j \\ 0, & \text{sinon} \end{cases}.$$

L'estimation du nombre de personnes, de la population de Suisse P, qui correspondent à la modalité j est alors donnée par :

$$\widehat{Y}_j = \sum_{i \in \tilde{R}} \bar{w}_i \cdot \delta_{ij},$$

et sa variance par la dernière formule du paragraphe 5.

Les Tableaux 10 (page 24) et 11 (page 25) donnent les estimations des totaux et leur écart-type

$$s = \sqrt{\widehat{Var}\left(\hat{Y}\right)}$$

obtenus pour la variable associée à la question Q193 : "Avez-vous été à un concert ou autre spectacle musical au cours des 12 derniers mois ? Et si oui à quelle fréquence ?". La personne interrogée pouvait choisir une des cinq réponses : "non", "oui, de une à trois fois", "oui, de quatre à six fois", "oui, de sept à douze fois" ou "oui, plus de douze fois".

Les totaux sont estimés par la procédure surveymeans de SAS, en utilisant les poids après calage $\bar{w}_i$. Les écarts-type sont calculés par la même procédure, en appliquant aux résidus $\hat{e}_i$ l'estimation de la variance pour un total dans le cas d'un échantillon stratifié, les tailles des strates étant données par les N_h du Tableau 4 (page 11).

La procédure surveymeans permet aussi d'estimer la proportion des personnes correspondant à la modalité j, ainsi que son écart-type, , en passant entre autre par la méthode des séries de Taylor, avec des formules analogues à celles utilisées pour les calculs précédents. Pour plus de détails concernant la procédure surveymeans voir [7].

Ces dernières estimations se trouvent dans le Tableau 12 (page 26).

TABLEAU 10 - Totaux et écarts-types.

Cellules	oui					
	1 à 3 fois		4 à 6 fois		7 à 12 fois	
	Totaux	s	Totaux	s	Totaux	s
Total	2'587'651	54'798	890'564	39'046	391'109	27'438
par sexe						
Homme	1'213'847	39'913	47'9797	29'381	209'812	21'295
Femme	1'373'804	37'572	41'0767	25'716	181'297	17'312
par classe d'âge						
15 *ans* - 19 *ans*	177'802	16'848	56'221	11'984	33'761	9'167
20 *ans* - 24 *ans*	189'187	18'693	60'474	12'824	25'992	7'433
25 *ans* - 34 *ans*	439'625	23'516	173'246	19'014	73'433	13'811
35 *ans* - 44 *ans*	534'095	23'289	164'915	14'997	48'620	8'729
45 *ans* - 54 *ans*	506'512	23'627	160'670	16'487	87'825	13'429
55 *ans* - 64 *ans*	363'652	19'126	150'235	14'274	48'268	8'446
65 *ans et plus*	376'777	18'784	124'802	12'428	73'210	9'767
par état civil						
marié(e), etc.	1'170'547	36'988	429'300	27'826	209'407	20'309
célibataire, etc.	1'417'103	40'446	461'265	27'374	181'702	18'459
par nationalité						
Suisse	1'894'757	44'322	654'681	32'037	298'562	23'055
Étranger	692'894	32'241	235'884	22'306	92'547	14'887

TABLEAU 11 - Totaux et écarts-types (suite).

Cellules	oui		non	
	plus de 12 fois		-	
	Totaux	s	Totaux	s
Total	281’957	23’765	2’042’875	50’614
par sexe				
Homme	149’609	17’792	978’993	37’427
Femme	132’348	15’759	1’063’881	34’160
par classe d’âge				
15 *ans* - 19 *ans*	30’902	9’263	136’794	15’740
20 *ans* - 24 *ans*	48’623	11’829	102’297	16’019
25 *ans* - 34 *ans*	59’900	10’880	229’333	19’748
35 *ans* - 44 *ans*	35’114	7’592	436’613	22’447
45 *ans* - 54 *ans*	40’937	8’370	297’536	21’130
55 *ans* - 64 *ans*	27’468	6’109	316’034	18’456
65 *ans et plus*	39’013	7’464	524’267	19’652
par état civil				
marié(e), etc.	185’086	19’955	796’036	32’084
célibataire, etc.	96’871	12’898	1’246’838	39’203
par nationalité				
Suisse	204’724	19’855	1’342’305	39’754
Étranger	77’233	13’059	700’569	31’346

TABLEAU 12 - Proportions et écarts-types.

Cellules	oui								non	
	1 à 3 fois		4 à 6 fois		7 à 12 fois		plus de 12 fois			
	%	s	%	s	%	s	%	s	%	s
Total	41.8%	0.9%	14.4%	0.6%	6.3%	0.4%	4.6%	0.4%	33.0%	0.8%
par sexe										
Homme	40.0%	1.3%	15.8%	1.0%	6.9%	0.7%	4.9%	0.6%	32.3%	1.2%
Femme	43.4%	1.2%	13.0%	0.8%	5.7%	0.5%	4.2%	0.5%	33.6%	1.1%
par classe d'âge										
15 *ans* - 19 *ans*	40.8%	3.9%	12.9%	2.8%	7.8%	2.1%	7.1%	2.1%	31.4%	3.6%
20 *ans* - 24 *ans*	44.4%	4.4%	14.2%	3.0%	6.1%	1.7%	11.4%	2.8%	24.0%	3.8%
25 *ans* - 34 *ans*	45.1%	2.4%	17.8%	1.9%	7.5%	1.4%	6.1%	1.1%	23.5%	2.0%
35 *ans* - 44 *ans*	43.8%	1.9%	13.5%	1.2%	4.0%	0.7%	2.9%	0.6%	35.8%	1.8%
45 *ans* - 54 *ans*	46.3%	2.2%	14.7%	1.5%	8.0%	1.2%	3.7%	0.8%	27.2%	1.9%
55 *ans* - 64 *ans*	40.2%	2.1%	16.6%	1.6%	5.3%	0.9%	3.0%	0.7%	34.9%	2.0%
65 *ans et plus*	33.1%	1.6%	11.0%	1.1%	6.4%	0.9%	3.4%	0.7%	46.0%	1.7%
par état civil										
marié(e), etc.	41.9%	1.3%	15.4%	1.0%	7.5%	0.7%	6.6%	0.7%	28.5%	1.1%
célibataire, etc.	41.6%	1.2%	13.6%	0.8%	5.3%	0.5%	2.8%	0.4%	36.6%	1.2%
par nationalité										
Suisse	43.1%	1.0%	14.9%	0.7%	6.8%	0.5%	4.7%	0.5%	30.5%	0.9%
Étranger	38.5%	1.8%	13.1%	1.2%	5.1%	0.8%	4.3%	0.7%	38.9%	1.7%

Références

[1] Andrade, B. et Salamin P.-A. (2012). Enquête sur la situation sociale et économique des étudinat·e·s des hautes écoles suisses 2009. Office fédéral de la statistique (OFS). Neuchâtel.

[2] Cochran, W.G. (1977). Sampling Techniques. John Wiley et Sons. New York.

[3] Deville, J.C. et Särndal, C.E. (1992). Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, 87, 376-382. Alexandria.

[4] Lohr, S.L. (2009). Sampling : Design and Analysis. Brooks/Cole Cengage Learning. Boston.

[5] Moeshler O., S. et Vanhooydonck, S. (2011). les pratiques culturelles en Suisse. Analyse approfondie - enquête 2008. OFS. Berne.

[6] Särndal, C.E. (2007). La méthode de calage dans la théorie et la pratique des enquêtes. Technique d'enquête, 33, 113-135. Ottawa.

[7] SAS/STAT® 9.1 (2004). User's Guide, 33, 4313-4362. SAS Publishing. Cary.

[8] Sautory, O. (2008). La macro CALMAR2. Redressement d'un échantillon par calage sur marges. Document de travail INSEE. Paris.

[9] Tillé, Y. (2001). Théorie des sondages. Dunod. Paris.

[10] Vanhooydonck, S. et Cuche, Y. (2007). Proposition de projet : Statistique sur les pratiques culturelles. Document interne OFS. Neuchâtel.

Methodenberichte der Sektion Statistische Methoden des BFS
Rapports de méthodes de la section méthodes statistiques de l'OFS
Methodology reports published by the FSO's Statistical Methods Section

Panchard, C. (2014). Enquête sur la participation aux activités culturelles 2008. Plan de sondage et estimations. Numéro de commande : 338-0073

Assoulin, D. (2014). Wertschöpfungsstatistik. Revision 2009 : Statistische Datenaufbereitung und Hochrechnung. Bestellnummer : 338-0072

Wilhelm, M. (2014). Echantillonnage boule de neige. La méthode de sondage déterminé par les répondants. Numéro de commande : 338-0071

Assoulin, D. (2013). Wertschöpfungsstatistik. Revision 2009 : Stichprobenrahmen und Stichprobenplan. Bestellnummer : 338-0070

Ferster, M. (2013). EVS I - Energieverbrauchsstatistik 2002 bis 2007 : Stichprobenplan und Hochrechnung. Bestellnummer : 338-0069

Assoulin, D. (2013). Zusatzerhebung für die landwirtschaftliche Betriebszählung 2010 : Stichprobenplan und Hochrechnung. Bestellnummer : 338-0068

Potterat, J. (2012). Use of conversion keys for NOGA2002-2008. Order number : 338-0067-05

Andrade, B., Salamin P.-A. (2012). Enquête sur la situation sociale et économique des étudiant-e-s des hautes écoles suisses 2009. Cadre de sondage, plan d'échantillonnage et méthodes d'estimation. Numéro de commande : 338-0066

Potterat, J., Panchard, C., Kilchmann, D. (2012). Umweltschutzausgaben der Unternehmen 2009 (UWSA2009). Stichprobenplan, Einsetzungen, Gewichtung und Schätzverfahren. Bestellnummer : 338-0065

Potterat, J. (2012). Benutzung der Umsteigeschlüssel NOGA 2002-2008. Bestellnummer : 338-0064

Potterat, J. (2011). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2009 (KNBB09). Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer : 338-0063

Kilchmann, D., Potterat, J., Genoud, S. (2011). Gütertransporterhebung 2008. Stichprobenplan, Datenaufbereitung, Gewichtung und Schätzverfahren. Bestellnummer : 338-0062

Graf, E. (2010). Enquête suisse sur la santé 2007. Plan d'échantillonnage, pondérations et analyses pondérées des données. Numéro de commande : 338-0061

Eichenberger P., Hulliger B., Potterat J. (2010). Describing the Anticipated Accuracy of the Swiss Population Survey. Order number : 338-0060

Graf, E. (2010). Étude empirique de l'attrition du Panel Suisse de Ménages : vers une caractérisation du profil du non-répondant. Numéro de commande : 338-0059

Graf, E. (2009). Weightings of the Swiss Household Panel : SHP_I wave 9, SHP_II wave 4, SHP_I et SHP_II combined. Order number : 338-0058

Graf, E. (2009). Pondérations du Panel Suisse de Ménages : PSM_I vague 9, PSM_II vague 4, PSM_I et PSM_II combinés. Numéro de commande : 338-0057-05

Qualité, L., Tillé, Y. (2009). Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée. Numéro de commande : 338-0056

Renaud, A., Panchard, C. et Potterat, J. (2008). Statistique de l'emploi. Révision 2007 : méthodes d'estimation. Numéro de commande : 338-0055

Graf, E. (2008). Pondérations du PSM. PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Numéro de commande : 338-0054

Andrade, B., Graf, M. (2008). Enquête suisse sur la structure des salaires 2006. Aspects méthodologiques du modèle des salaires "Salarium". Numéro de commande : 338-0053

Renaud, A. (2008). Statistique de l'emploi. Révision 2007 : cadre de sondage et échantillonage. Numéro de commande : 338-0052

Graf, E. (2008). Pondérations du SILC pilote. SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Numéro de commande : 338-0051

Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. Bestellnummer : 338-0050

Renaud, A. (2008). Technologies de l'information et de la communication. Estimations sur la base de la statistique de la valeur ajoutée. Numéro de commande : 338-0049

Assoulin, D. (2007). Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen. Bestellnummer : 338-0048

Kilchmann, D. (2007). Beherbergungsstatistik Campingplätze. Stichprobenrahmen und Schätzverfahren 2005/06. Bestellnummer : 338-0047

Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. Bestellnummer : 338-0046

Ferrez, J., Graf, M. (2007). Enquête suisse sur la structure des salaires. Programmes R pour l'intervalle de confiance de la médiane. Numéro de commande : 338-0045

Renaud, A. (2007). Harmonisation de la scolarité obligatoire en Suisse (HarmoS). Design général de l'enquête et échantillon des écoles. Numéro de commande : 338-0044

Potterat, J. (2007). Betriebszählung 2005. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer : 338-0043

Hulliger, B. (2006). Umweltschutzausgaben der Unternehmen 2003, Stichprobenplan, Datenaufbereitung und Schätzverfahren. Bestellnummer : 338-0042

Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro de commande : 338-0041

Salamin, P.-A. (2006). Statistique de l'aide sociale dans le domaine de l'asile. Plan de sondage et extrapolations pour l'enquête pilote 2005. Numéro de commande : 338-0040

Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro de commande : 338-0039

Graf, M. (2006). Swiss Earnings Structure Survey 2002-2004. Compositional data in a stratified two-stage sample : Analysis and precision assessment of wage components. Order number : 338-0038

Potterat, J. (2006). Pensionskassenstatistik 2004. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer : 338-0037

Potterat, J. (2006). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2004. Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer : 338-0036

Kilchmann, D. (2006). Vierteljährliche Wohnbaustatistik. Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005. Bestellnummer : 338-0035

Kilchmann, D. (2006). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2004. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer : 338-0034

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 2. Bestellnummer : 338-0033

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 1. Bestellnummer : 338-0032

Graf, M., Matei, A. (2005). Enquête suisse sur la structure des salaires 2002. La précision du salaire brut standardisé médian. Numéro de commande : 338-0031

Graf, E., Renfer, J.-P. (2005). Enquête suisse sur la santé 2002. Plan d'échantillonnage, pondération et estimation de la précision. Numéro de commande : 338-0030

Potterat, J. (2005). Mietpreis-Strukturerhebung 2003. Gewichtung und Schätzverfahren. Bestellnummer : 338-0029

Potterat, J. (2005). Landwirtschaftliche Betriebszählung 2003. Schätzverfahren für die Zusatzerhebung. Bestellnummer : 338-0028

Renaud, A. (2004). Coverage estimation for the Swiss population census 2000. Estimation methodology and results. Order number : 338-0027

Kilchmann, D. (2004). Revision des Schweizerischen Lohnindex. Schätzmethoden der Lohnindices und deren Varianzschätzer. Bestellnummer : 338-0026

Graf, M. (2004). Enquête suisse sur la structure des salaires 2002. Plan d'échantillonnage et extrapolation pour le secteur privé. Numéro de commande : 338-0025

Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro de commande 338-0024

Renaud, A., Potterat, J. (2004). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sous-couverture (P-sample) et qualité du cadre de sondage des bâtiments. Numéro de commande : 338-0023

Graf, M. (2004). Fusion de données. Etude de faisabilité. Numéro de commande : 338-0022

Potterat, J. (2003). Mietpreis-Strukturerhebung 2003. Entwicklung des Stichprobenplans und Ziehung der Stichprobe. Bestellnummer : 338-0021

Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. Bestellnummer : 338-0020.

Renaud, A. (2003). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sur-couverture (E-sample). Numéro de commande : 338-0019

Hulliger, B. (2003). Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Erhebung über F+E in der schweizerischen Privatwirtschaft 2000. Bestellnummer : 338-0018

Renfer, J.-P. (2003). Enquête 2000 sur la recherche et le développement dans l'économie privée en Suisse. Plan d'échantillonnage. Numéro de commande : 338-0017

Potterat, J. (2003). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe. Schätzverfahren. Bestellnummer : 338-0016

Graf, M., Matei, A. (2003). Stratégie de choix des modèles de désaisonnalisation. Application aux séries de l'emploi total. Numéro de commande : 338-0015

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Methoden für die Datenbereinigung. Bestellnummer : 338-0014

Renaud, A. (2002). Programme international pour le suivi des acquis des é lèves (PISA). Plans d'échantillonnage pour PISA 2000 en Suisse. Numéro de commande : 338-0013

Renfer, J.-P. (2002). Enquête 2001 sur les coûts et l'utilité de la formation des apprentis du point de vue des établissements. Plan d' échantillonnage. Numéro de commande : 338-0012

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Stichprobenplan und Schätzverfahren für die provisorischen Ergebnisse. Bestellnummer : 338-0011

Graf, M. (2002). Enquête suisse sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé. Numéro de commande : 338-0010

Renaud, A., Eichenberger P. (2002). Estimation de la couverture du recensement de la population de l'an 2000. Procédure d'enquête et plan d'échantillonnage de l'enquête de couverture. Numéro de commande : 338-0009

Kilchmann, D., Hulliger, B. (2002). Stichprobenplan für die Obstbaumzählung 2001. Bestellnummer : 338-0008

Graf, M. (2002). Passage du concept établissement au concept entreprise. Numéro de commande : 338-0007

Salamin, P.A. (2001). La technique de la double enquête pour la statistique du transport routier de marchandise. Numéro de commande : 338-0006

Peters, R., Renfer, J.-P. et Hulliger, B. (2001). Statistique de la valeur ajoutée 1997-1998. Procédure d'extrapolation des données. Numéro de commande : 338-0005

Potterat, J., Hulliger, B. (2001). Schätzung der Sägereiproduktion mit der Sägerei-Erhebung PAUL. Bestellnummer : 338-0004

Graf, M. (2001). Désaisonnalisation. Aspects méthodologiques et application à la statistique de l'emploi. Numéro de commande : 338-0003

Hüsler, J., Müller, S. (2001). Schlussbericht Betriebszählung 1995 (BZ 95), Mehrfach imputierte Umsatzzahlen. Bestellnummer : 338-0002

Renaud, A. (2001). Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Numéro de commande : 338-0001

Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen : Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. Bestellnummer : 338-0000

Programme des publications de l'OFS

En sa qualité de service central de statistique de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public.

L'information statistique est diffusée par domaine (cf. verso de la première page de couverture); elle emprunte diverses voies:

Moyen de diffusion	*Contact*
Service de renseignements individuels	058 463 60 11 info@bfs.admin.ch
L'OFS sur Internet	www.statistique.admin.ch
Communiqués de presse: information rapide concernant les résultats les plus récents	www.news-stat.admin.ch
Publications: information approfondie	058 463 60 60 order@bfs.admin.ch
Données interactives (banques de données, accessibles en ligne)	www.stattab.bfs.admin.ch

Informations sur les divers moyens de diffusion sur Internet à l'adresse www.statistique.admin.ch → Services → Les publications de Statistique suisse

Rapports de méthodes de la section méthodes statistiques

Les rapports de méthodes décrivent les méthodes mathématiques et statistiques à la base des résultats et des analyses de la statistique publique. Ils présentent également l'évaluation et le développement de nouvelles méthodes en vue d'une application future. Ces publications visent d'une part à documenter les méthodes utilisées ou envisagées dans un souci de transparence et de rigueur scientifique, et d'autre part à favoriser la collaboration avec le monde scientifique et universitaire.

Les résultats numériques présentés dans les rapports de méthodes illustrent les concepts mathématiques décrits, mais ne sont pas des résultats officiels des enquêtes concernées. De même, les méthodes réellement appliquées peuvent différer légèrement de celles décrites dans ces rapports.

Les rapports de méthodes sont disponibles sous forme électronique sur le site internet de l'OFS.

Ce rapport documente le plan de sondage, l'allocation et le tirage de l'échantillon, ainsi que les méthodes d'estimation utilisées pour la statistique concernant la participation aux activités culturelles. Cette statistique se base sur une enquête auprès des ménages résidents en Suisse. Le plan d'échantillonage est conçu de manière à pouvoir mettre en évidence les différences culturelles relatives aux régions linguistiques de la Suisse. Ces dernières constituent donc les strates du tirage et l'allocation choisie tient compte de la précision et de la non-réponse au niveau de ces strates. L'échantillon se compose de 7'423 personnes pour lesquelles on obtient 4'346 répondants. La pondération de ces répondants comporte une correction de la non-réponse et un calage, de telle sorte qu'elle permet des extrapolations de l'échantillon à la population de la Suisse. Ces extrapolations sont accompagnées d'une estimation de variance.

N° de commande
338-0073

Commandes
Tél.: 058 463 60 60
Fax: 058 463 60 61
E-mail: order@bfs.admin.ch

Prix
gratuit

ISBN 978-3-303-00517-0