



HESTA-Mon

Revenu sur l'hébergement dans le secteur hôtelier

Pondération et estimation de variance

EXPERIMENTAL STATISTICS



Neuchâtel, 2024

Éditeur:	Office fédéral de la statistique (OFS)	Concept de mise en page:	section PUB
Renseignements:	info-tour@bfs.admin.ch, tél. +41 58 463 61 97	Téléchargement:	www.statistique.ch
Rédaction:	section METH, OFS	Copyright:	OFS, Neuchâtel 2024
Contenu:	Raphaël Jauslin, OFS; Rémy Mariétan, OFS	La reproduction est autorisée, sauf à des fins commerciales, si la source est mentionnée	
Domaine:	00 Bases statistiques des généralités		
Langue du texte original:	Français		

Table des matières

1	Introduction	1
2	Population et variables	1
2.1	Notations et objectifs	2
3	Traitement de la non-réponse	3
3.1	Modélisation du mécanisme de non-réponse	4
4	Calage	10
4.1	Fonction de calage	12
5	Variance	15
5.1	Variance d'un estimateur mensuel	15
5.1.1	Total	15
5.1.2	Ratio	15
5.2	Variance d'estimateur sur des périodes	16
5.2.1	Total	16
5.2.2	Ratio	20
5.3	Analyse des variances	20
6	Exemples de résultats	22
7	Limitations	24
	Références	25
A	Annexe - tel reminder et response channel	26
B	Annexe - Évaluation des estimateurs	27
C	Annexe - Évaluation du calage	28

1 Introduction

La statistique de l'hébergement touristique (HESTA) est réalisée à l'échelle nationale au niveau du secteur hôtelier auprès d'environ 4 500 établissements. Son but est de fournir des informations sur le nombre d'établissements, de chambres, de lits et de nuitées. Dans ce cadre, la section Méthodes statistiques (METH) a été sollicitée pour proposer une méthode d'estimation du revenu par nuitée ainsi que du revenu total des établissements sur des périodes données.

La section Tourisme (TOUR) s'intéresse à différentes périodes comme l'année civile, la saison d'été (de mai à octobre) et la saison d'hiver (de novembre à avril). Les chiffres doivent pouvoir être ventilés par nombre d'étoiles et zone touristique agrégée.

L'objectif est de fournir un estimateur du revenu moyen par nuitée et par mois ainsi que du revenu total par mois. Afin de réduire les risques de biais dus à la non-réponse et d'augmenter la précision de l'estimation, un modèle de non-réponse ainsi qu'une procédure de calage sont proposés dans ce document.

Pour construire des estimateurs saisonniers ou annuels, nous avons choisi de combiner des estimateurs mensuels. Cette approche, dont l'avantage principal est l'utilisation de toutes les données disponibles, requiert une estimation non triviale de la variance des estimations. Ce document explique ainsi également comment cette variance a été estimée.

2 Population et variables

L'enquête exhaustive porte sur environ 4 500 établissements chaque mois. Le cadre est constitué de l'ensemble des établissements suisses. Plusieurs variables sont disponibles dans le cadre et sont utilisées pour modéliser le mécanisme de non-réponse et pour le calage. Les variables disponibles sur tout le cadre qui seront prépondérantes sont les suivantes :

<code>instance_name</code>	Le mois et l'année de la réponse de l'établissement.
<code>category_id</code>	Nombre d'étoiles de l'établissement. (aucune information, 0, 1, 2, 3, 4, 5).
<code>room_count</code>	Nombre de chambres de l'établissement.
<code>bed_count</code>	Nombre de lits de l'établissement.
<code>st_zone</code>	Zone touristique (montagne 1, campagne 2, ville 3, métropole 4)

Les variables d'intérêt utilisées dans le cadre de ce projet qui sont connues uniquement sur les répondants sont les suivantes :

<code>nights_total</code>	Nombre de nuitées total.
<code>avg_income_night</code>	Revenu moyen par nuitée, renseigné par les établissements.
<code>open_days</code>	Nombre de jours ouverts.
<code>response_channel</code>	Canal de réponse.
<code>response_created_stamp</code>	Date et heure d'enregistrement de la réponse.

Pour les établissements considérés comme répondants dans ce projet (c'est-à-dire qui

renseignent `nights_total` et `avg_income_night`), le revenu total mensuel se déduit donc en multipliant `avg_income_night` par `nights_total`. L'enquête a comme but principal de relever le nombre de nuitées par pays de provenance et la question sur le revenu suit ensuite dans le questionnaire. Donc, il n'y a pas de répondants qui renseignent leurs revenus sans renseigner leurs nuitées. À l'inverse, il y a des établissements qui ne renseignent pas leurs revenus mais qui ont renseigné leurs nuitées. Le mécanisme de non-réponse peut ainsi être considéré en deux phases.

2.1 Notations et objectifs

Pour la suite de ce document, U_τ définira le cadre du mois τ et N_τ le nombre d'unités dans U_τ . Dans un souci de faciliter la lecture et puisque ce document traite souvent d'un seul mois spécifique, la notation est souvent simplifiée. La notation du mois sera ainsi généralement omise, $U_\tau \equiv U$ et $N_\tau \equiv N$. Lorsque plusieurs mois sont traités simultanément, cela est précisé dans le document. Soit y_k le revenu mensuel d'un établissement $k \in U$ et z_k le nombre de nuitées qu'il a renseigné, le but est d'estimer le revenu total sur le mois

$$Y = \sum_{k \in U} y_k$$

et le nombre de nuitées total

$$Z = \sum_{k \in U} z_k.$$

On s'intéresse au revenu moyen par nuitée. L'objectif est donc d'estimer le ratio suivant :

$$Q = \frac{Y}{Z}.$$

Chaque total Y et Z sera estimé en utilisant un estimateur par calage

$$\hat{Y} = \sum_{k \in R} w_k y_k, \quad \hat{Z} = \sum_{k \in R} w_k z_k,$$

où R correspond à l'ensemble des répondants. Le ratio est estimé simplement en faisant le ratio des estimateurs calés :

$$\hat{Q} = \frac{\hat{Y}}{\hat{Z}}.$$

La section TOUR aimerait pouvoir ventiler les estimateurs par nombre d'étoiles et/ou zones touristiques, ceux-ci étant intégrés par les traitements présentés dans la suite du document. La section 3 décrit la modélisation du mécanisme de non-réponse et la section 4 le calage avec chaque fois leurs procédures pour choisir les variables pertinentes. Ensuite, en section 5, la variance est développée dans le but de produire une mesure de précision des estimateurs. Des exemples de résultats sont alors présentés en section 6. Enfin, la section 7 présente les limitations de l'approche choisie.

TABLE 1 – Taux de réponse mensuel, où N représente le nombre d'établissements dans le cadre du mois correspondant, n_1 représente le nombre d'établissements ayant renseigné le nombre de nuitées, n_2 représente le nombre d'établissements ayant renseigné le revenu moyen par nuitée et le nombre de nuitées.

Année-mois	N	n_1	n_2	n_1/N	n_2/N
2022-07	4 501	4 384	2 475	0.97	0.55
2022-08	4 491	4 376	2 478	0.97	0.55
2022-09	4 490	4 366	2 490	0.97	0.56
2022-10	4 486	4 353	2 551	0.97	0.57
2022-11	4 483	4 339	2 968	0.97	0.66
2022-12	4 481	4 343	2 687	0.97	0.60
2023-01	4 515	4 463	2 877	0.99	0.64
2023-02	4 510	4 449	2 871	0.99	0.64
2023-03	4 506	4 438	3 203	0.98	0.71
2023-04	4 500	4 420	3 275	0.98	0.73
2023-05	4 508	4 435	3 352	0.98	0.74
2023-06	4 505	4 416	3 211	0.98	0.71

3 Traitement de la non-réponse

Le mécanisme de non-réponse est séparé en deux parties. Le premier ensemble R_1 représente les établissements qui ont renseigné le nombre de nuitées pour un mois τ . Le deuxième ensemble R_2 représente les établissements ayant renseigné le revenu moyen par nuitée. Le deuxième ensemble est imbriqué dans le premier $R_2 \subset R_1$ par la nature de l'enquête. La table 1 montre les taux de réponse des deux niveaux, le taux de réponse du deuxième niveau augmente au fil des mois grâce des procédures de rappel mises en place par la section à cette période. Les taux de réponse du deuxième niveau sont nettement plus bas que ceux du premier niveau.

Depuis mars 2023 la section TOUR a mis en place une procédure de rappel pour les établissements 5 étoiles ou ceux possédant plus de 150 lits afin d'augmenter le taux de réponse du deuxième niveau dans ces catégories. La table 2 montre les taux de réponse ventilés par nombre d'étoiles pour quatre mois en 2023. La figure 1 montre l'évolution des taux de réponse du deuxième niveau sur la période 2023-01 à 2023-12.

TABLE 2 – Taux de réponse mensuel et par nombre d'étoiles (les catégories 1 et 2 étoiles ont été regroupées et la catégorie 100 représente la catégorie sans information), N est le nombre d'établissements dans le cadre du mois correspondant, n_1 le nombre d'établissements ayant renseigné le nombre de nuitées, n_2 le nombre d'établissements ayant renseigné le revenu moyen par nuitée.

Année-mois	Nombre d'étoiles	N	n_1	n_2	n_1/N	n_2/N
2023-03	0	308	304	209	0.99	0.68
2023-03	3	1 082	1 076	743	0.99	0.69
2023-03	4	528	527	395	1.00	0.75
2023-03	5	112	112	99	1.00	0.88
2023-03	12	140	137	84	0.98	0.60
2023-03	100	2 336	2 282	1 673	0.98	0.72
2023-04	0	309	303	213	0.98	0.69

(Continue sur la page suivante...)

TABLE 2 – Taux de réponse mensuel et par nombre d'étoiles (les catégories 1 et 2 étoiles ont été regroupées et la catégorie 100 représente la catégorie sans information), N est le nombre d'établissements dans le cadre du mois correspondant, n_1 le nombre d'établissements ayant renseigné le nombre de nuitées, n_2 le nombre d'établissements ayant renseigné le revenu moyen par nuitée. (suite)

Année-mois	Nombre d'étoiles	N	n_1	n_2	n_1/N	n_2/N
2023-04	3	1 082	1 072	762	0.99	0.70
2023-04	4	528	528	415	1.00	0.79
2023-04	5	111	111	97	1.00	0.87
2023-04	12	141	138	86	0.98	0.61
2023-04	100	2 329	2 268	1 702	0.97	0.73
2023-05	0	309	305	231	0.99	0.75
2023-05	3	1 082	1 078	819	1.00	0.76
2023-05	4	529	529	436	1.00	0.82
2023-05	5	111	111	102	1.00	0.92
2023-05	12	141	138	89	0.98	0.63
2023-05	100	2 336	2 274	1 675	0.97	0.72
2023-06	0	310	305	235	0.98	0.76
2023-06	3	1 083	1 076	767	0.99	0.71
2023-06	4	529	529	420	1.00	0.79
2023-06	5	111	111	99	1.00	0.89
2023-06	12	141	136	76	0.96	0.54
2023-06	100	2 331	2 259	1 614	0.97	0.69

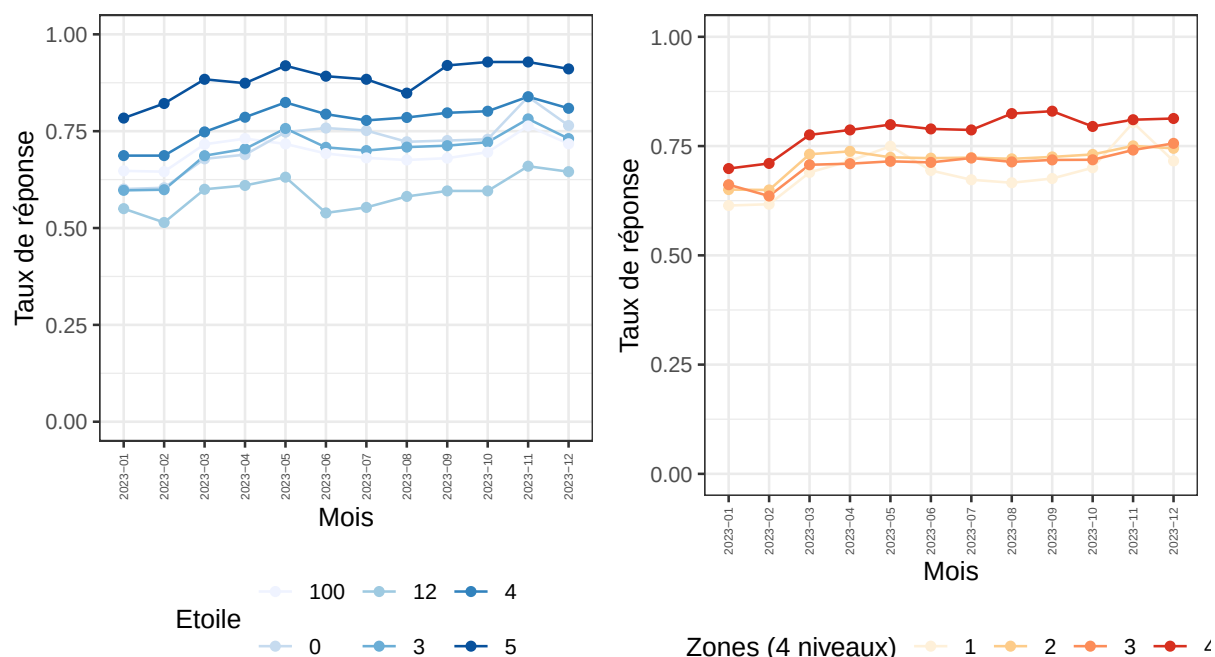


FIGURE 1 – Taux de réponse du deuxième niveau par nombre d'étoiles et zones.

3.1 Modélisation du mécanisme de non-réponse

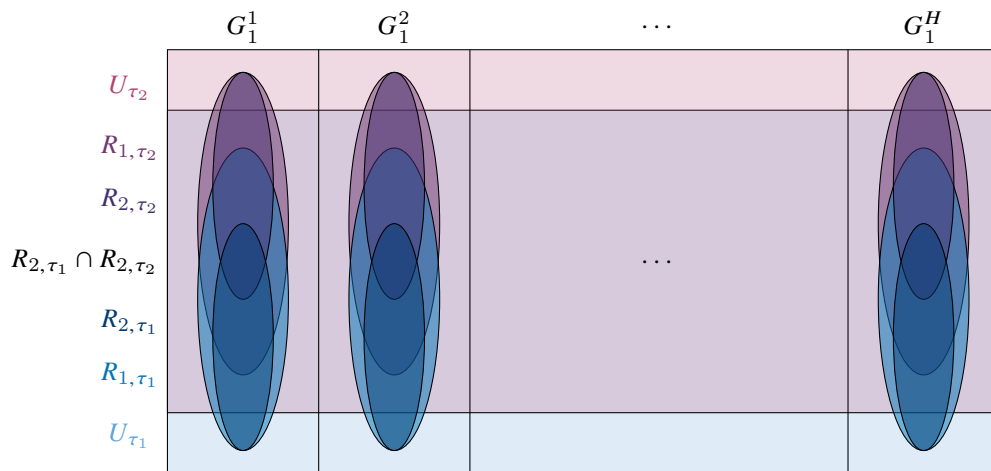
Le mécanisme de non-réponse est modélisé en utilisant des groupes homogènes de réponse avec l'hypothèse que le mécanisme de non-réponse est similaire à l'intérieur

des groupes. La probabilité de réponse sera donc unique au sein d'un même groupe. Ces groupes sont formés sur la base de variables auxiliaires qui expliquent partiellement la probabilité de réponse et qui sont également corrélées avec les variables d'intérêt. En effet, certaines variables peuvent expliquer le mécanisme de non-réponse tout en ayant peu de corrélation avec les variables d'intérêts. Par exemple, les variables `response_channel` et `tel_reminder` expliquent bien le mécanisme de non-réponse, mais ne sont que très peu corrélées avec le revenu moyen ou le revenu total (cf. annexe A). L'ajout d'une telle variable va certes améliorer la modélisation de la non-réponse, mais peut avoir un impact négatif sur les estimateurs de nos variables d'intérêt (instabilité, groupes homogènes moins pertinents pour les variables d'intérêt...).

Les groupes homogènes sont d'abord créés pour le premier niveau puis affinés pour le deuxième niveau. Soient $G_1^h \subset U$ les groupes homogènes du premier niveau $h \in \{1, 2, \dots, H\}$ et $G_2^\ell \subset R_1$ les groupes homogènes du deuxième niveau $\ell \in \{1, 2, \dots, L\}$. Soient $R_1^h \subset U$ les répondants des groupes homogènes du premier niveau $h \in \{1, 2, \dots, H\}$ et $R_2^\ell \subset R_1$ les répondants des groupes homogènes du deuxième niveau $\ell \in \{1, 2, \dots, L\}$. Les tables 3 et 4 montrent les croisements des variables précédemment citées pour former les groupes. La figure 2 illustre la population et sa subdivision en groupes homogènes ainsi que les différents ensembles présentés précédemment entre deux mois τ_1 et τ_2 . La figure 3 montre les évolutions des taux de réponse dans chaque groupe du premier et du deuxième niveau. Deux procédures du logiciel SAS ont été utilisées afin d'obtenir ces groupes homogènes. Ces procédures sont `proc HPSPLIT` et `proc logisitc`. Nous n'avons cependant pas directement utilisé les sorties des procédures telles quelles, mais nous les avons utilisées comme soutien à la décision.

Grâce aux répondants du premier niveau R_1 , l'information qui renseigne si l'établissement était actif durant le mois est disponible pour traiter les non-répondants du deuxième niveau. Un établissement peut par exemple être en vacances ou décider pour une autre raison de fermer un mois. Deux points sont importants à souligner pour la formation des groupes G_2^ℓ .

FIGURE 2 – Illustration des ensembles $G_1, U_{\tau_1}, U_{\tau_2}, R_{1,\tau_1}, R_{1,\tau_2}, R_{2,\tau_1}, R_{2,\tau_2}$ pour deux mois donnés τ_1 et τ_2 . Le deuxième niveau de groupes homogènes G_2 est volontairement omis afin de ne pas surcharger l'image. Il pourrait cependant être représenté par une portion de la différence entre l'ellipse R_{1,τ_1} et R_{2,τ_1} .



1. Les établissements qui ont renseigné le fait qu'ils étaient en vacances ou qui n'ont aucun jour d'ouverture seront mis dans un groupe à part appelé les non-actifs. Ils reçoivent une probabilité de réponse pour le deuxième niveau égale à 1.
2. Afin de ne pas faire exploser le nombre de groupes G_2^ℓ , i.e. garder L raisonnablement petit, la variable du nombre de nuitées a été catégorisée pour essayer d'incorporer cette information dans les groupes homogènes. Cependant, une analyse du biais a montré que le gain n'est pas suffisant par rapport à l'augmentation du nombre de groupes. Autrement dit, le nombre de groupes L devient grand et le biais ne se réduit que de très peu (voir annexe B pour plus de détails). Ainsi il a été décidé de ne pas utiliser cette variable pour la formation des groupes homogènes du deuxième niveau.

TABLE 3 – Groupes homogènes utilisés pour modéliser le mécanisme de non-réponse aux nuitées (premier niveau). Taille min. est égal au plus petit nombre d'établissements par mois au sein de chaque groupe et Taux est égal au taux de réponse moyen par mois. Période 2023-01 à 2023-12.

Étoiles	Nombres de lits	Zones touristiques	Taille min.	Taux
5	-	-	111	1.00
4	1-70	1-2	107	1.00
4	71-150	1-2	129	1.00
4	1-70	3-4	41	1.00
4	71-150	3-4	79	1.00
4	151+	-	170	1.00
3	1-70	1-2	620	0.99
3	71-150	1-2	157	0.99
3	1-70	3-4	147	0.99
3	71-150	3-4	80	0.99
3	151+	-	75	1.00
100-0-1-2	1-70	1-2	2 093	0.97
100-0-1-2	71-150	1-2	167	0.99
100-0-1-2	1-70	3-4	312	0.96
100-0-1-2	71-150	3-4	58	1.00
100-0-1-2	151+	-	101	0.99

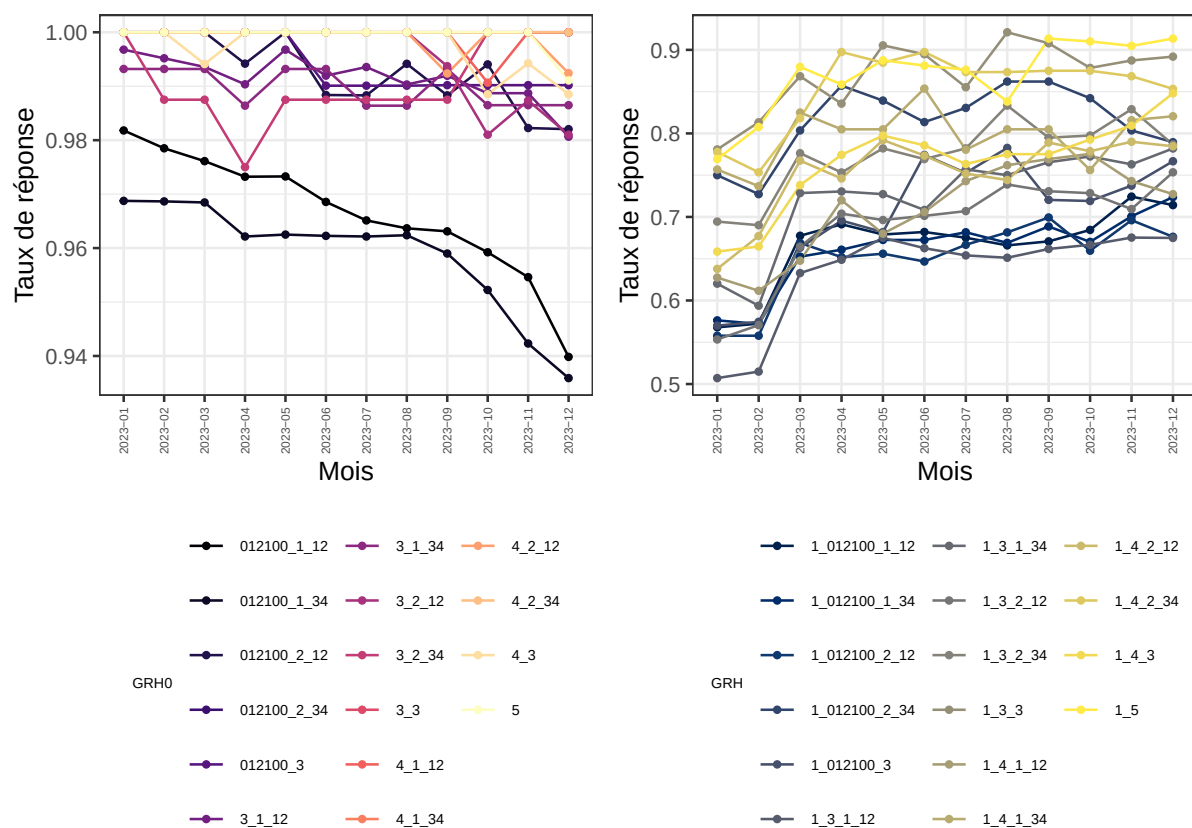


FIGURE 3 – Évolution des taux de réponse par groupe homogène du premier et deuxième niveau. Le premier chiffre représente le nombre d'étoiles (012100, 3, 4, 5), le deuxième chiffre la catégorie du nombre de lits (1, 2, 3) et finalement la catégorie des zones (12, 34). Pour les groupes homogènes du deuxième niveau le code commence par le nombre 1 qui représente uniquement les actifs.

TABLE 4 – Groupes homogènes formés pour modéliser le mécanisme de non-réponse du deuxième niveau (celui qui renseigne le revenu moyen). Taille min. est égal au plus petit nombre d'établissements par mois se trouvant dans chaque groupe et Taux est égal au taux de réponse moyen par mois. Période 2023-01 à 2023-12.

Jours ouverts > 0	Étoiles	Nombre de lits	Zone touristique	Taille min.	Taux
0	-	-	-	220	1.00
1	5	-	-	80	0.87
1	4	1-70	1-2	70	0.71
1	4	71-150	1-2	100	0.75
1	4	1-70	3-4	37	0.80
1	4	71-150	3-4	72	0.85
1	4	151+	-	152	0.76
1	3	1-70	1-2	425	0.64
1	3	71-150	1-2	117	0.69
1	3	1-70	3-4	128	0.72
1	3	71-150	3-4	71	0.77
1	3	151+	-	71	0.87
1	100-0-1-2	1-70	1-2	1 375	0.67
1	100-0-1-2	71-150	1-2	102	0.65
1	100-0-1-2	1-70	3-4	257	0.66
1	100-0-1-2	71-150	3-4	55	0.81
1	100-0-1-2	151+	-	80	0.70

Soit $k \in U$, les probabilités modélisées sont les suivantes :

$$p_{1,k}^h = P(k \in R_1^h, k \in G_1^h) \text{ et } p_{2,k}^\ell = P(k \in R_2^\ell, k \in G_2^\ell).$$

Soient $|G_1^h|$ le nombre d'unités dans le groupe G_1^h , $|R_1^h|$ le nombre de répondants du premier niveau dans le groupe G_1^h , $|G_2^\ell|$ le nombre d'unités dans le groupe G_2^ℓ et $|R_2^\ell|$ le nombre de répondants du deuxième niveau dans le groupe G_2^ℓ . Notons que le nombre de répondants du premier niveau dans le groupe G_2^ℓ vaut exactement $|G_2^\ell|$ car les non-actifs sont considérés comme répondants au deuxième niveau i.e. dont on apprend leur non-activité seulement lors de l'enquête en cours. Ceci vient de l'imbrication des groupes homogènes du premier et du deuxième niveau. Ces probabilités sont estimées par :

$$\hat{p}_{1,k}^h = \frac{|R_1^h|}{|G_1^h|}, k \in G_1^h \subset U \text{ et } \hat{p}_{2,k}^\ell = \frac{|R_2^\ell|}{|G_2^\ell|}, k \in G_2^\ell \subset S_1 \text{ où } h \in \{1, 2, \dots, H\}, \ell \in \{1, 2, \dots, L\}$$

Lors de la formation des groupes pour le premier niveau de réponse, seuls le nombre d'étoiles, la zone touristique et le nombre de lits sont utilisés. L'hypothèse admise est que les établissements n'ayant pas d'information ou ayant zéro, une ou deux étoiles possèdent un mécanisme de non-réponse similaire entre eux et des revenus moyens proches. De plus, les établissements à cinq étoiles forment un groupe à eux seuls.

Pour un mois, on estime donc la probabilité de renseigner le revenu moyen de l'établissement k ,

$$\hat{\theta}_k = \hat{p}_{1,k}^h \hat{p}_{2,k}^\ell.$$

Ainsi, les poids corrigeant la non-réponse valent pour $k \in R_2$ tel que $k \in G_2^\ell \subset G_1^h$

$$\frac{1}{\hat{\theta}_k} = \frac{1}{\hat{p}_{1,k}^h \hat{p}_{2,k}^\ell}.$$

Ces poids seront optimisés dans la section suivante par l'étape du calage.

4 Calage

Dans le but d'augmenter la qualité de l'estimateur, un calage des échantillons de chaque mois sur leur cadre mensuel U_τ a été appliqué. Dans cette section nous nous concentrons sur le calage d'un seul mois sur son cadre que nous notons U . Soit x_j une variable de calage, le total sur le cadre mensuel est donné par l'équation :

$$X_j = \sum_{k \in U} x_{j,k}.$$

Les poids $1/\hat{\theta}_k$ seront modifiés de telle manière que les poids finaux obtenus, notés w_k , soient proches des poids avant calage et que les totaux X_j des variables auxiliaires connues sur le cadre, soient égaux à ceux obtenus en appliquant les poids w_k à R_2 . Mathématiquement cela revient à écrire :

$$X_j = \sum_{k \in U} x_{j,k} = \sum_{k \in R_2} w_k x_{j,k} = \hat{X}_j.$$

Comme expliqué dans [Tillé \(2020\)](#), un principe de parcimonie doit être appliqué et il est important d'utiliser des variables auxiliaires qui sont corrélées avec les variables d'intérêts. Les variables de calage sont choisies pour améliorer la précision des estimations. Il faut toutefois que les nombres d'unités dans les groupes créés par les variables auxiliaires catégorielles soient suffisamment grand. Ci-dessous la liste des variables de calage qui ont été choisies :

- Zone touristique (1,2,3,4)
- nombre d'étoiles (0,1,2,3,4,5, pas d'information)
- Indicatrice du nombre de lits > 150
- Nombre d'étoiles (0,1,2,3,4,5, pas d'information) x nombre de lits
- Zone touristique (1,2,3,4) x nombre de lits
- Zone touristique agrégé (12,34) x nombre d'étoiles
- Zone touristique agrégé (12,34) x nombre d'étoiles x nombre de lits

La table [5](#) montre le nombre d'observations dans les groupes formés par les variables de calage. Pour plus de détails sur les variables de calage et les marges pondérées voir annexe [C](#).

TABLE 5 – Table du nombre d'observations au sein des marges de calage pour le mois de juin 2023. Seulement une sélection de ceux-ci sont affichées ici.

Mois	Zone	Etoile	Big bed	Etoile x zone	Nombre d'observations
2023-06				0_12	192
2023-06				0_34	43
2023-06				100_12	1 384
2023-06				100_34	230
2023-06				12_12	51
2023-06				12_34	25
2023-06				3_12	572
2023-06				3_34	195
2023-06				4_12	241
2023-06				4_34	179
2023-06				5_12	57
2023-06				5_34	42
2023-06			0		2 869
2023-06			1		342
2023-06		0			235
2023-06		3			767
2023-06		4			420
2023-06		5			99
2023-06		12			76
2023-06		100			1 614
2023-06	1				1 641
2023-06	2				856
2023-06	3				377
2023-06	4				337

TABLE 6 – Résumé de la distribution des poids avant et après calage.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Avant calage	1	1.262	1.49	1.408	1.535	1.978
Après calage	1	1.222	1.43	1.408	1.569	2.635

4.1 Fonction de calage

On effectue un calage linéaire tronqué à l'aide du logiciel SAS et de la procédure CALMAR ([Le Guennec & Sautory, 2002](#)) sur des marges ajustées afin de garantir des poids après calage supérieurs à 1 (voir le rapport de méthode [Chevalier & Qualité \(2021\)](#) pour plus de détails). Mathématiquement, on a

$$\sum_{k \in R_2} w_k x_{j,k} - \sum_{k \in R_2} x_{j,k} = \sum_{k \in U} x_{j,k} - \sum_{k \in R_2} x_{j,k},$$

qui donne après simplification :

$$\sum_{k \in R_2} (w_k - 1) x_{j,k} = X_j^* = \sum_{k \in U} x_{j,k} - \sum_{k \in R_2} x_{j,k},$$

où X_j^* représente les marges du cadre moins les marges de l'échantillon, donc les marges ajustées. Si on note $w'_k = w_k - 1$ des poids calés intermédiaires, cela revient à une équation de calage classique. Les bornes inférieure et supérieure pour le calage linéaire tronqué sont égales à (0,50). Il suffit alors d'ajouter 1 aux poids w'_k pour obtenir les poids après calages w_k . Les figures 4 et 5 et la table 6 montrent la distribution des poids avant et après calage. Les poids après calage gardent une distribution proche des poids initiaux et aucun poids extrême n'est observé. Les pics dans les graphiques de la figure 5 proviennent de l'hypothèse d'une probabilité de réponse unique au sein des groupes homogènes et du nombre élevé d'unités dans certains de ces groupes.

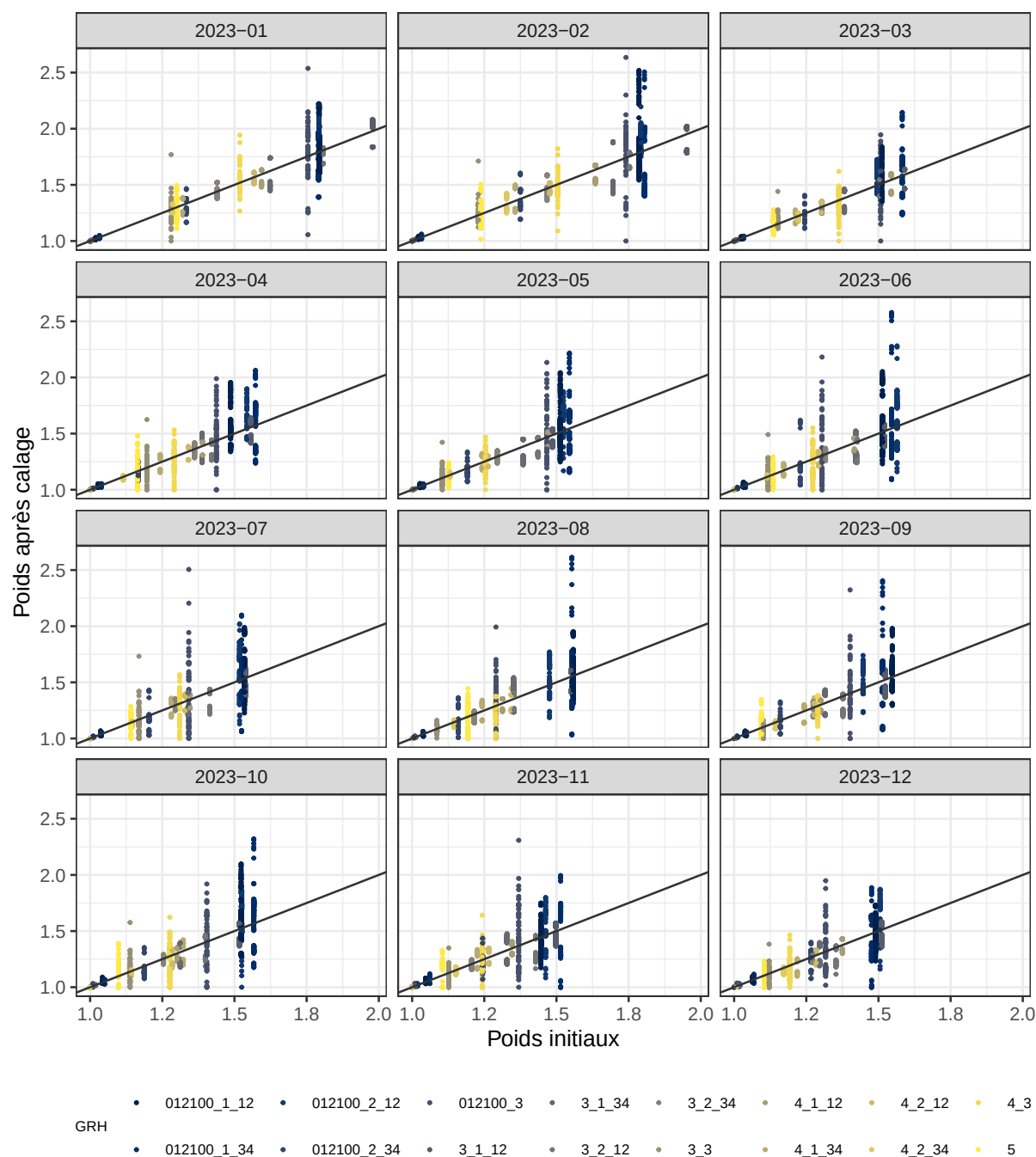


FIGURE 4 – Distribution des poids avant et après calage sur le premier semestre 2023.

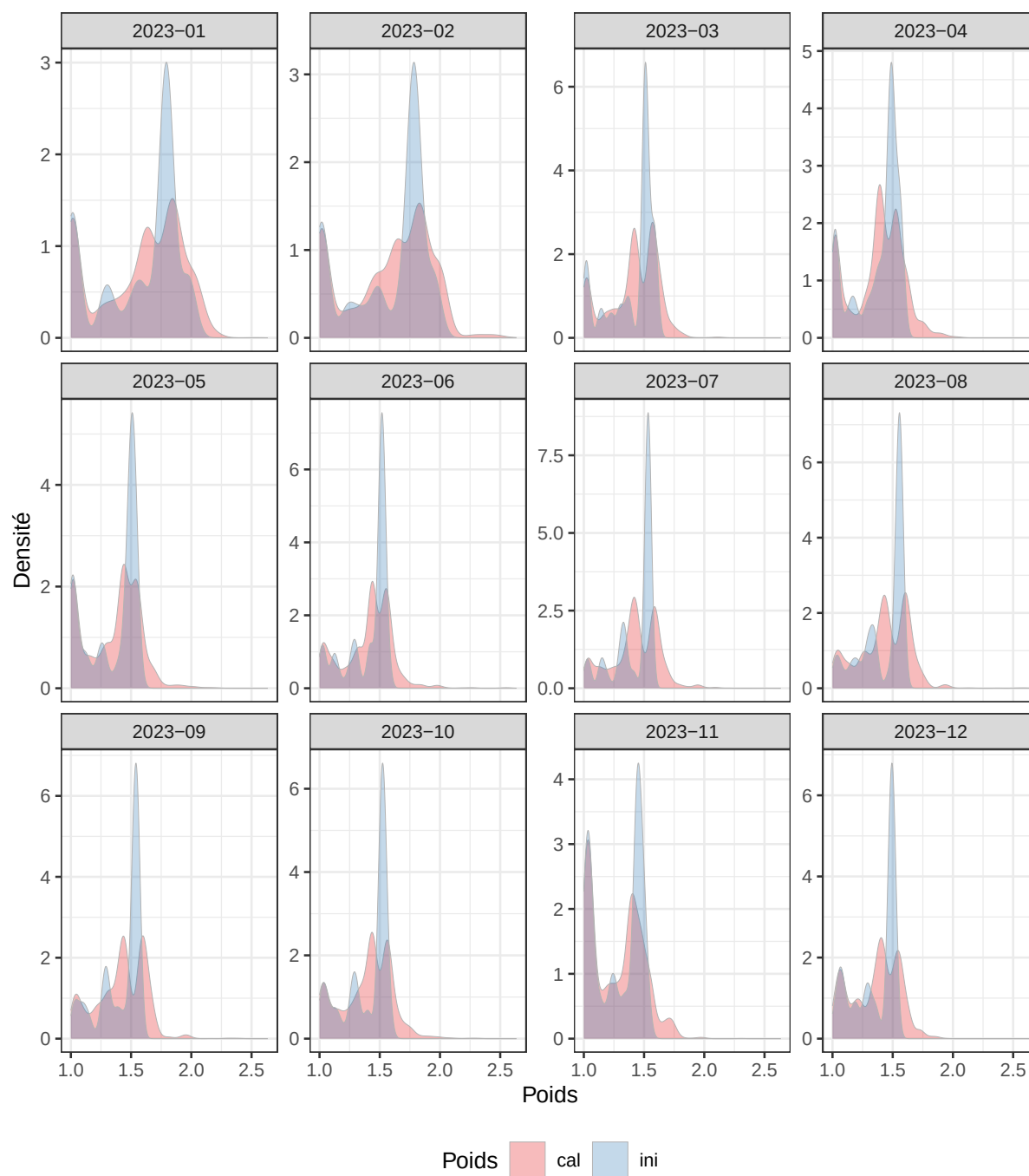


FIGURE 5 – Densité des poids avant et après calage du premier semestre 2023. L'abréviation 'ini' correspond aux poids avant calage et 'cal' aux poids après calage.

5 Variance

Dans cette section, le développement des estimateurs de variance est présenté, d'abord pour les estimations mensuelles, puis pour celles effectuées sur un regroupement de mois.

5.1 Variance d'un estimateur mensuel

5.1.1 Total

À l'aide des poids après calage, les estimateurs du revenu total mensuel et du nombre de nuitées sont donnés par :

$$\widehat{Y} = \sum_{k \in R_2} w_k y_k, \quad \widehat{Z} = \sum_{k \in R_2} w_k z_k$$

Grâce aux travaux de [Deville & Särndal \(1992\)](#), la variance de l'estimateur calé peut-être approché par l'utilisation des résidus d'une régression multiple. Cette variance peut s'écrire de la manière suivante :

$$\text{var}(\widehat{Y}) \simeq \text{var} \left(\sum_{k \in R_2} \frac{1}{\widehat{\theta}_k} E_k \right) \quad (1)$$

où $1/\widehat{\theta}_k$ est considéré comme un poids fixe non-aléatoire et $E_k = y_k - \mathbf{x}_k^\top \mathbf{B}$ sont les résidus de la régression de la variable d'intérêt y_k sur les variables de calage. Cette régression étant ajustée sur la population U , ces résidus E_k ne sont pas disponibles en pratique. [Deville & Särndal \(1992\)](#) proposent donc d'utiliser l'échantillon pour estimer cette variance. Dans le cadre de la théorie de l'échantillonnage au sein de populations finies, on peut utiliser soit les poids de calage w_k soit les poids avant calage d_k pour obtenir un estimateur de la variance. Une discussion théorique sur le choix des poids avec des références aux articles y relatifs peut se trouver dans [Tillé \(2020\)](#), pages 365-366. L'estimateur de variance que nous utilisons s'écrit :

$$\begin{aligned} \widehat{\text{var}}(\widehat{Y}) &= \widehat{\text{var}} \left(\sum_{k \in R_2} \frac{1}{\widehat{\theta}_k} E_k \right) \\ &= \sum_{k \in R_2} \frac{1 - \widehat{\theta}_k}{\widehat{\theta}_k^2} e_k^2, \end{aligned} \quad (2)$$

où e_k est le résidu de la régression calculé sur l'échantillon. L'égalité s'obtient sous l'hypothèse que le mécanisme de non-réponse peut être vu comme un tirage de Poisson indépendant (voir [Deville & Särndal \(1992\)](#) pour plus de détails).

5.1.2 Ratio

L'estimateur du ratio est donné par l'équation :

$$\widehat{Q} = \frac{\widehat{Y}}{\widehat{Z}}.$$

Le principe de linéarisation d'un estimateur développé par [Deville \(1999\)](#) est utilisé pour donner une approximation de la variance :

$$\widehat{\text{var}}(\widehat{Q}) = \widehat{\text{var}}\left(\sum_{k \in R_2} w_k \widehat{l}_k\right),$$

où

$$\widehat{l}_k = \frac{1}{\widehat{Z}} \left(y_k - \frac{\widehat{Y}}{\widehat{Z}} z_k \right),$$

est considéré comme fixe. Il suffit alors d'utiliser la formule (2) en utilisant la quantité $\widehat{l}_k$ au lieu de la variable d'intérêt y_k pour estimer les résidus (voir [Tillé \(2020\)](#), Chapitre 15, pour plus de détails).

5.2 Variance d'estimateur sur des périodes

5.2.1 Total

Pour l'estimation sur une période, par exemple la saison d'été (mai-octobre), la saison d'hiver (novembre-mai) ou l'année civile (janvier-décembre), un regroupement de mois $\mathcal{T}$ est utilisé. L'estimation du total est donnée par l'équation :

$$\widehat{Y}_{\mathcal{T}} = \sum_{\tau \in \mathcal{T}} \widehat{Y}_{\tau} = \sum_{\tau \in \mathcal{T}} \sum_{k \in R_{2,\tau}} w_k y_k.$$

On suppose que les établissements participent indépendamment les uns des autres, mais qu'il y a une dépendance dans la participation d'un même établissement à différents mois. De plus, en supposant que le mécanisme de non-réponse est un tirage de Poisson indépendant, alors la variance de l'estimateur sur le groupement de mois $\mathcal{T}$ peut s'approximer de la manière suivante :

$$\begin{aligned}
\text{var}\left(\sum_{\tau \in \mathcal{T}} \widehat{Y}_\tau\right) &= \sum_{\tau \in \mathcal{T}} \text{var}\left(\widehat{Y}_\tau\right) + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \text{cov}\left(\widehat{Y}_{\tau_1}, \widehat{Y}_{\tau_2}\right) \\
&= \sum_{\tau \in \mathcal{T}} \text{var}\left(\sum_{k \in R_{2,\tau}} w_{\tau,k} y_{\tau,k}\right) \\
&\quad + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \text{cov}\left(\sum_{k \in R_{2,\tau_1}} w_{\tau_1,k} y_{\tau_1,k}, \sum_{k \in R_{2,\tau_2}} w_{\tau_2,k} y_{\tau_2,k}\right) \\
&\stackrel{*1}{\approx} \sum_{\tau \in \mathcal{T}} \text{var}\left(\sum_{k \in R_{2,\tau}} \frac{1}{\widehat{\theta}_{\tau,k}} E_{\tau,k}\right) \\
&\quad + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \text{cov}\left(\sum_{k \in R_{2,\tau_1}} \frac{1}{\widehat{\theta}_{\tau_1,k}} E_{\tau_1,k}, \sum_{k \in R_{2,\tau_2}} \frac{1}{\widehat{\theta}_{\tau_2,k}} E_{\tau_2,k}\right) \\
&= \sum_{\tau \in \mathcal{T}} \text{var}\left(\sum_{k \in U_\tau} \frac{1}{\widehat{\theta}_{\tau,k}} E_{\tau,k} 1_{k \in R_{2,\tau}}\right) \\
&\quad + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \text{cov}\left(\sum_{k \in U_{\tau_1}} \frac{1}{\widehat{\theta}_{\tau_1,k}} E_{\tau_1,k} 1_{k \in R_{2,\tau_1}}, \sum_{k \in U_{\tau_2}} \frac{1}{\widehat{\theta}_{\tau_2,k}} E_{\tau_2,k} 1_{k \in R_{2,\tau_2}}\right) \\
&\stackrel{*2}{\approx} \sum_{\tau \in \mathcal{T}} \sum_{k \in U_\tau} \frac{1}{\theta_{\tau,k}^2} E_{\tau,k}^2 \text{var}(1_{k \in R_{2,\tau}}) \\
&\quad + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \sum_{k \in U_{\tau_1} \cap U_{\tau_2}} \frac{1}{\theta_{\tau_1,k}} E_{\tau_1,k} \frac{1}{\theta_{\tau_2,k}} E_{\tau_2,k} \text{cov}(1_{k \in R_{2,\tau_1}}, 1_{k \in R_{2,\tau_2}}) \\
&= \sum_{\tau \in \mathcal{T}} \sum_{k \in U_\tau} \frac{p_{\tau,k} - p_{\tau,k}^2}{\theta_{\tau,k}^2} E_{\tau,k}^2 \\
&\quad + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \sum_{k \in U_{\tau_1} \cap U_{\tau_2}} \frac{p_{\tau_1 \tau_2,k} - p_{\tau_1,k} p_{\tau_2,k}}{\theta_{\tau_1,k} \theta_{\tau_2,k}} E_{\tau_1,k} E_{\tau_2,k}
\end{aligned} \tag{3}$$

où $p_{\tau,k}$ représente la probabilité de réponse inconnue et $1_{k \in R_2}$ représente la variable indicatrice valant 1 si l'unité k appartient à l'échantillon des répondants R_2 et 0 sinon. L'approximation $*1$ s'obtient par la linéarisation d'un estimateur calé (voir (1)) et l'approximation $*2$ néglige l'impact de l'échantillon aléatoire sur l'estimation des θ (normalement faible avec un estimateur calé voir Kim & Kim (2007)). On peut alors estimer la variance avec la formule suivante :

$$\begin{aligned}
\widehat{\text{var}}\left(\sum_{\tau \in \mathcal{T}} \widehat{Y}_\tau\right) &= \sum_{\tau \in \mathcal{T}} \sum_{k \in R_2} \frac{1 - \widehat{p}_{\tau,k|U_\tau}}{\widehat{\theta}_{\tau,k}^2} e_k^2 \\
&\quad + \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \sum_{k \in R_{2,\tau_1} \cap R_{2,\tau_2}} \frac{\widehat{p}_{\tau_1 \tau_2,k|U_{\tau_1} \cap U_{\tau_2}}}{\widehat{p}_{\tau_1 \tau_2,k|U_{\tau_1} \cap U_{\tau_2}} \widehat{\theta}_{\tau_1,k} \widehat{\theta}_{\tau_2,k}} e_{\tau_1,k} e_{\tau_2,k} \\
&\quad - \sum_{\substack{\tau_1, \tau_2 \in \mathcal{T} \\ \tau_1 \neq \tau_2}} \sum_{k \in R_{2,\tau_1} \cap R_{2,\tau_2}} \frac{\widehat{p}_{\tau_1,k|U_{\tau_1} \cap U_{\tau_2}} \widehat{p}_{\tau_2,k|U_{\tau_1} \cap U_{\tau_2}}}{\widehat{p}_{\tau_1 \tau_2,k|U_{\tau_1} \cap U_{\tau_2}} \widehat{\theta}_{\tau_1,k} \widehat{\theta}_{\tau_2,k}} e_{\tau_1,k} e_{\tau_2,k}
\end{aligned} \tag{4}$$

Pour estimer les variances ci-dessus, il semble approprié de poser $\hat{\theta}_{\tau,k} = \hat{p}_{\tau,k|U_{\tau}}$. Les premiers termes de la formule se simplifient en estimateur de la variance mensuelle classique.

Bien que $\hat{p}_{\tau_1,k|U_{\tau_1} \cap U_{\tau_2}}$ et $\hat{\theta}_{\tau_1,k} = \hat{p}_{\tau_1,k|U_{\tau_1}}$ estiment la même probabilité $p_{\tau_1,k}$, il peut être préférable de les garder différents pour exploiter les informations sur les cadres et pour estimer la covariance de manière plus cohérente en respectant les inégalités ci-dessous :

$$\begin{aligned}\hat{p}_{\tau_1\tau_2,k|U_{\tau_1} \cap U_{\tau_2}} &\leq \hat{p}_{\tau_1,k|U_{\tau_1} \cap U_{\tau_2}}, \\ \hat{p}_{\tau_1\tau_2,k|U_{\tau_1} \cap U_{\tau_2}} &\leq \hat{p}_{\tau_2,k|U_{\tau_1} \cap U_{\tau_2}}.\end{aligned}$$

Pour chaque paire de mois, il faut estimer les quantités suivantes : $\hat{p}_{\tau_1,k|U_{\tau_1} \cap U_{\tau_2}}$, $\hat{p}_{\tau_2,k|U_{\tau_1} \cap U_{\tau_2}}$ et $\hat{p}_{\tau_1\tau_2,k|U_{\tau_1} \cap U_{\tau_2}}$.

Soit $G_{1,\tau_1}^h, G_{1,\tau_2}^h \subset U$, $h \in \{1, \dots, H\}$ les groupes homogènes du premier niveau et $G_{2,\tau_1}^{\ell}, G_{2,\tau_2}^{\ell} \subset S_1$, $\ell \in \{1, \dots, L\}$ les groupes homogènes du deuxième niveau pour les mois τ_1 et τ_2 . Il s'agit alors d'estimer la probabilité qu'une unité réponde au premier et au deuxième niveau sachant qu'elle appartient à l'intersection des deux mois, i.e. $k \in U_{\tau_1} \cap U_{\tau_2}$. Soient $R_{1,\tau_1}, R_{2,\tau_1}, R_{1,\tau_2}$ et R_{2,τ_2} les ensembles des répondants des mois τ_1 et τ_2 .

Méthode 1

Il s'agit tout d'abord d'estimer la probabilité :

$$\begin{aligned}p_{\tau_1,k|U_{\tau_1} \cap U_{\tau_2}} &= P(k \in R_{2,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}) \\ &= P(k \in R_{2,\tau_1} | k \in R_{1,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}) \times P(k \in R_{1,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}).\end{aligned}$$

Ce qui donne :

$$\begin{aligned}\hat{p}_{\tau_1,k|U_{\tau_1} \cap U_{\tau_2}} &= \hat{P}(k \in R_{2,\tau_1} | k \in R_{1,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}) \times \hat{P}(k \in R_{1,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}) \\ &= \frac{|G_{2,\tau_1}^{\ell} \cap R_{2,\tau_1}|}{|G_{2,\tau_1}^{\ell} \cap R_{1,\tau_1}|} \frac{|G_{1,\tau_1}^h \cap R_{1,\tau_1}|}{|G_{1,\tau_1}^h|}, \quad h \in \{1, \dots, H\}, \ell \in \{1, \dots, L\}.\end{aligned} \quad (5)$$

La probabilité $p_{\tau_1\tau_2,k|U_{\tau_1} \cap U_{\tau_2}}$ se calcule naturellement en utilisant les mêmes principes de probabilités conditionnelles :

$$\begin{aligned}p_{\tau_1\tau_2,k|U_{\tau_1} \cap U_{\tau_2}} &= P(k \in R_{2,\tau_1} \cap R_{2,\tau_2} | k \in U_{\tau_1} \cap U_{\tau_2}) \\ &= P(k \in R_{2,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}) \\ &\quad \times P(k \in R_{2,\tau_2} | k \in R_{2,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}) \\ &= P(k \in R_{2,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}) \\ &\quad \times P(k \in R_{1,\tau_2} | k \in R_{2,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}) \\ &\quad \times P(k \in R_{2,\tau_2} | k \in R_{1,\tau_2} \cap R_{2,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}).\end{aligned}$$

On obtient l'estimation suivante :

$$\begin{aligned}
\widehat{P}_{\tau_1\tau_2,k|U_{\tau_1}\cap U_{\tau_2}} &= \widehat{P}(k \in R_{2,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}) \\
&\quad \times \widehat{P}(k \in R_{1,\tau_2} | k \in R_{2,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}) \\
&\quad \times \widehat{P}(k \in R_{2,\tau_2} | k \in R_{1,\tau_2} \cap R_{2,\tau_1} \cap U_{\tau_1} \cap U_{\tau_2}) \\
&= \frac{|G_{2,\tau_1}^\ell \cap R_{2,\tau_1}| |G_{1,\tau_1}^h \cap R_{1,\tau_1}|}{|G_{2,\tau_1}^\ell \cap R_{1,\tau_1}| |G_{1,\tau_1}^h|} \\
&\quad \times \frac{|G_{1,\tau_2}^h \cap R_{1,\tau_2} \cap R_{2,\tau_1}|}{|G_{1,\tau_2}^h \cap R_{2,\tau_1}|} \\
&\quad \times \frac{|G_{2,\tau_2}^\ell \cap R_{2,\tau_1} \cap R_{2,\tau_2}|}{|G_{2,\tau_2}^\ell \cap R_{2,\tau_1} \cap R_{1,\tau_2}|},
\end{aligned} \tag{6}$$

où $h \in \{1, \dots, H\}$, $\ell \in \{1, \dots, L\}$. Finalement pour estimer $p_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}}$, le même procédé que pour (5) peut être appliqué. Ainsi, pour estimer

$$\begin{aligned}
p_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}} &= P(k \in R_{2,\tau_2} | k \in U_{\tau_1} \cap U_{\tau_2}) \\
&= P(k \in R_{2,\tau_2} | k \in R_{1,\tau_2} \cap U_{\tau_1} \cap U_{\tau_2}) \times P(k \in R_{1,\tau_2} | k \in U_{\tau_1} \cap U_{\tau_2}),
\end{aligned}$$

on calcule

$$\begin{aligned}
\widehat{p}_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}} &= \widehat{P}(k \in R_{2,\tau_2} | k \in R_{1,\tau_2} \cap U_{\tau_1} \cap U_{\tau_2}) \times \widehat{P}(k \in R_{1,\tau_2} | k \in U_{\tau_1} \cap U_{\tau_2}) \\
&= \frac{|G_{2,\tau_2}^\ell \cap R_{2,\tau_2}| |G_{1,\tau_2}^h \cap R_{1,\tau_2}|}{|G_{2,\tau_2}^\ell \cap R_{1,\tau_2}| |G_{1,\tau_2}^h|},
\end{aligned} \tag{7}$$

où $h \in \{1, \dots, H\}$, $\ell \in \{1, \dots, L\}$. Notons qu'il est tout à fait possible de calculer (7) puis de l'utiliser afin de calculer (6). Cela ne donnera pas le même résultat, car nous utiliserions d'autres probabilités conditionnelles et cela donnera une estimation différente. A priori, ces estimations devraient être proches.

Méthode 2

Le deuxième traitement possible estime tout d'abord la probabilité de réponse d'un premier mois comme dans les équations (5) et (7). Dans l'approche qui suit nous supposons que nous calculons d'abord $\widehat{p}_{\tau_1,k|U_{\tau_1}\cap U_{\tau_2}}$. Ensuite, en utilisant la formule des probabilités totales sur l'ensemble des répondants/non-répondants du premier niveau au deuxième mois ainsi que de l'estimation de la probabilité conjointe (6), on peut proposer une autre approche pour estimer la probabilité $p_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}}$. Les équations (5) et (6) sont donc utilisées afin d'estimer $p_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}}$, ce qui donne

$$\begin{aligned}
p_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}} &= P(k \in R_{2,\tau_2} | k \in U_{\tau_1} \cap U_{\tau_2}) \\
&= P(k \in R_{2,\tau_2} \cap R_{2,\tau_1} | k \in U_{\tau_1} \cap U_{\tau_2}) + P(k \in R_{2,\tau_2} \cap R_{2,\tau_1}^c | k \in U_{\tau_1} \cap U_{\tau_2})
\end{aligned}$$

que l'on peut estimer en utilisant la formule

$$\begin{aligned}
\widehat{p}_{\tau_2,k|U_{\tau_1}\cap U_{\tau_2}} &= \widehat{p}_{\tau_1\tau_2,k|U_{\tau_1}\cap U_{\tau_2}} + \widehat{P}(k \in R_{2,\tau_1}^c | k \in U_{\tau_1} \cap U_{\tau_2}) \\
&\quad \times \widehat{P}(k \in R_{1,\tau_2} | k \in R_{2,\tau_1}^c \cap U_{\tau_1} \cap U_{\tau_2}) \\
&\quad \times \widehat{P}(k \in R_{2,\tau_2} | k \in R_{1,\tau_2} \cap R_{2,\tau_1}^c \cap U_{\tau_1} \cap U_{\tau_2}).
\end{aligned} \tag{8}$$

Cette approche peut être réalisée dans l'autre sens en calculant d'abord $\widehat{p}_{\tau_2,k|U_{\tau_1} \cap U_{\tau_2}}$ et ne donnera pas le même résultat. Ces différentes méthodes d'estimations de variances sont programmées et comparées dans le programme fourni à la section TOUR. Pour des estimations mensuelles, ces variances sont proches. En effet, les estimations de covariance peuvent différer entre les deux approches sans pour autant impacter fortement l'estimation de variance sur une période. Cependant, en fonction des ventilations, il n'est pas exclu d'observer des différences plus importantes. Si les estimations sont très différentes, alors de plus amples analyses devraient être effectuées pour choisir l'approche plus fiable.

5.2.2 Ratio

Le but est d'estimer des ratios de totaux sur un groupement de mois. Par exemple l'estimation du revenu moyen durant l'été, l'hiver ou l'année civile. L'estimateur a la forme suivante :

$$\widehat{Q}_{\mathcal{T}} = \frac{\widehat{Y}_{\mathcal{T}}}{\widehat{Z}_{\mathcal{T}}} = \frac{\sum_{\tau \in \mathcal{T}} \widehat{Y}_{\tau}}{\sum_{\tau \in \mathcal{T}} \widehat{Z}_{\tau}}.$$

Comme pour les ratios mensuels, le principe de linéarisation de l'estimateur [Deville \(1999\)](#) peut être utilisé. L'estimation de la variance est donc donnée par l'équation suivante :

$$\widehat{\text{var}}(\widehat{Q}_{\mathcal{T}}) = \widehat{\text{var}}\left(\sum_{\tau \in \mathcal{T}} \left\{ \sum_{k \in R_{2,\tau}} w_k \widehat{l}_{\tau,k} \right\}\right),$$

où

$$\widehat{l}_{\tau,k} = \frac{1}{\widehat{Z}_{\mathcal{T}}} \left(y_{\tau,k} - \frac{\widehat{Y}_{\mathcal{T}}}{\widehat{Z}_{\mathcal{T}}} z_{\tau,k} \right),$$

est considéré comme fixe.

5.3 Analyse des variances

La variance s'obtient en combinant de nombreuses covariances qui peuvent être calculées différemment. Dans cette section, l'impact de l'estimation des covariances sur la variance estimée est analysé. Les analyses observent le revenu moyen des établissements sur la période de 2023-01 à 2023-12.

Lorsque 12 estimateurs sont additionnés (1 année), il existe une matrice de covariance $V \in \mathbb{R}^{12} \times \mathbb{R}^{12}$, où $V_{i,j} = \widehat{\text{cov}}(\widehat{Y}_i, \widehat{Y}_j)$ représente la covariance estimée entre les estimateurs du $i^{\text{ème}}$ et du $j^{\text{ème}}$ mois de la période analysée. Pour l'estimation à l'aide de la méthode 1 en section 5.2.1, la figure 6 montre que les termes de variance entre les éléments diagonaux et les autres sont du même ordre de grandeur même si les covariances sont en règle générale plus basses. Sur la base de ces observations, on peut conclure qu'une estimation de la variance d'un estimateur cumulé sur 12 mois ne peut pas négliger les termes de covariance en faisant ainsi l'hypothèse d'indépendance entre les mois.

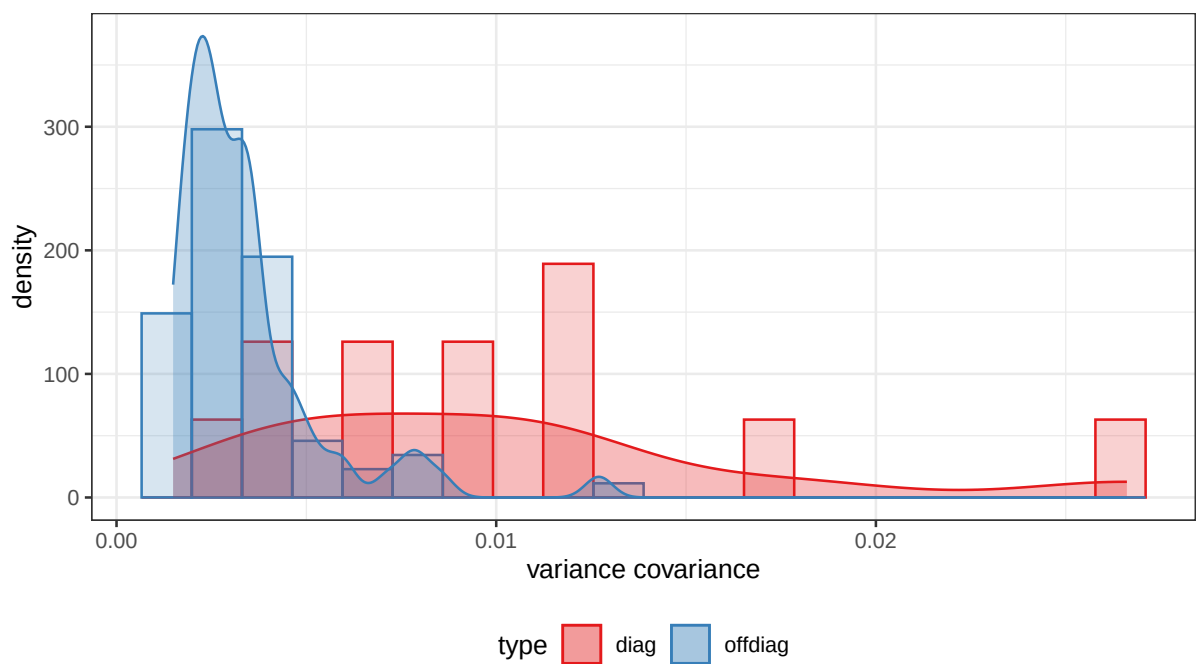


FIGURE 6 – Distribution des termes de covariance pour l'estimation du ratio sur une période d'une année (2023-01 à 2023-12). La méthode utilisée pour calculer les termes est la méthode 1 présentée en section [5.2.1](#)

6 Exemples de résultats

Dans cette section, quelques exemples de résultats sont présentés. Les figures 7 et la table 7 montrent les résultats obtenus pour l'estimation des revenus moyens pour la période 2023-01 2023-12. Selon le nombre d'étoiles, et d'un mois à l'autre, les variances sont passablement différentes. Les mois d'hiver semblent avoir une variance plus élevée avec des estimations qui sont elles aussi plus hautes.

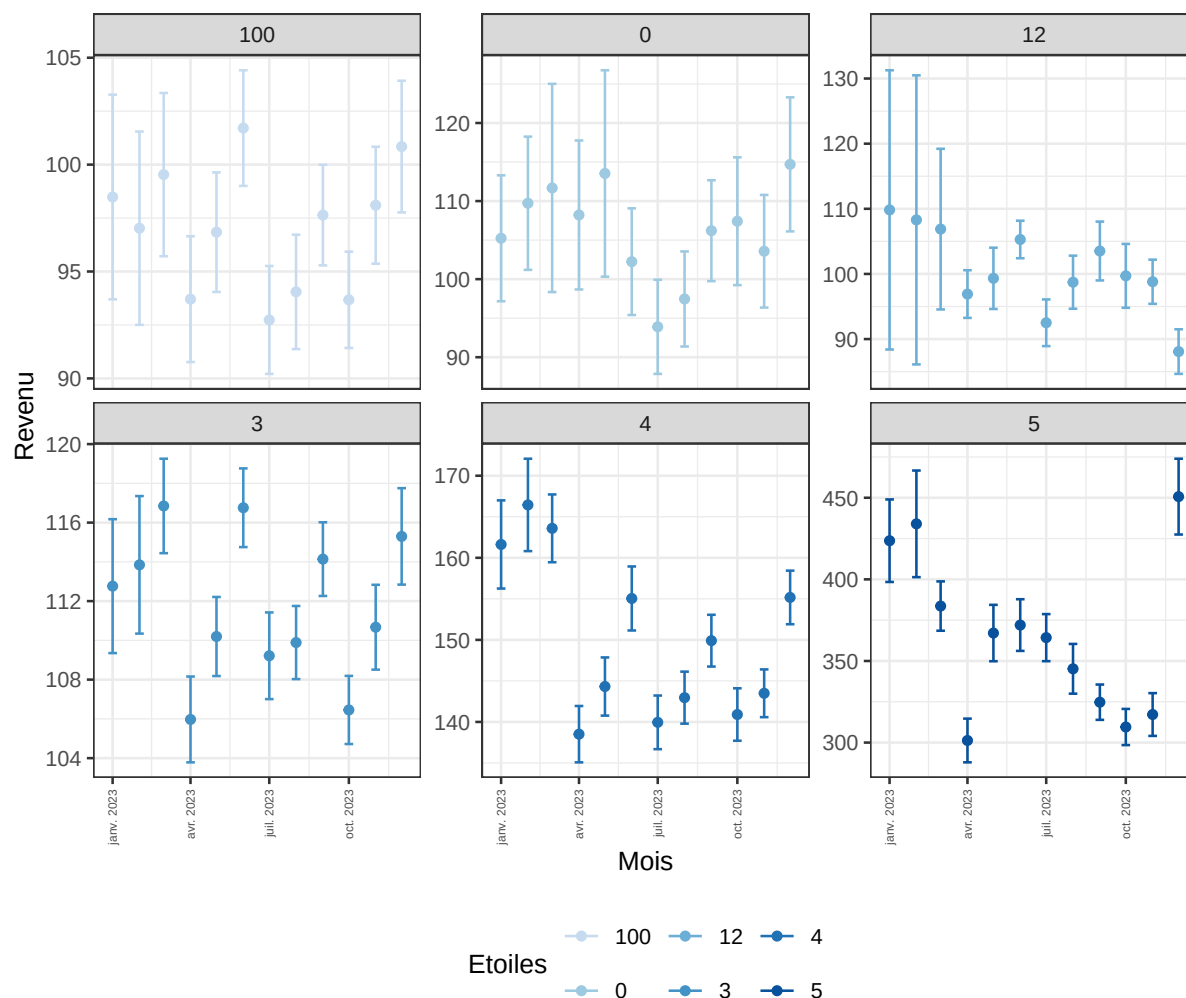


FIGURE 7 – Estimation des ratios correspondants aux revenus moyens avec intervalle de confiance pour la période allant de janvier 2023 à décembre 2023.

TABLE 7 – Estimation des ratios correspondants aux revenus moyens pour la période allant de janvier 2023 à décembre 2023.

Période	Etoile	Estimateur	Ecart-type	cv
2023-01	0	105	4.1	0.039
2023-02	0	110	4.4	0.040
2023-03	0	112	6.8	0.061
2023-04	0	108	4.9	0.045
2023-05	0	114	6.7	0.059
2023-06	0	102	3.5	0.034
2023-07	0	94	3.1	0.033
2023-08	0	98	3.1	0.032
2023-09	0	106	3.3	0.031
2023-10	0	107	4.2	0.039
2023-11	0	104	3.7	0.036
2023-12	0	115	4.4	0.038
2023-01 2023-12	0	105	2.8	0.027
2023-01	100	98	2.4	0.025
2023-02	100	97	2.3	0.024
2023-03	100	100	1.9	0.020
2023-04	100	94	1.5	0.016
2023-05	100	97	1.4	0.015
2023-06	100	102	1.4	0.014
2023-07	100	93	1.3	0.014
2023-08	100	94	1.4	0.015
2023-09	100	98	1.2	0.012
2023-10	100	94	1.1	0.012
2023-11	100	98	1.4	0.014
2023-12	100	101	1.6	0.016
2023-01 2023-12	100	97	1.1	0.011
2023-01	12	110	10.9	0.100
2023-02	12	108	11.3	0.105
2023-03	12	107	6.3	0.059
2023-04	12	97	1.9	0.019
2023-05	12	99	2.4	0.024
2023-06	12	105	1.5	0.014
2023-07	12	92	1.8	0.020
2023-08	12	99	2.1	0.021
2023-09	12	104	2.3	0.022
2023-10	12	100	2.5	0.025
2023-11	12	99	1.7	0.018
2023-12	12	88	1.7	0.020
2023-01 2023-12	12	100	2.2	0.022
2023-01	3	113	1.7	0.015
2023-02	3	114	1.8	0.016
2023-03	3	117	1.2	0.011
2023-04	3	106	1.1	0.011
2023-05	3	110	1.0	0.009
2023-06	3	117	1.0	0.009
2023-07	3	109	1.1	0.010
2023-08	3	110	1.0	0.009
2023-09	3	114	1.0	0.008
2023-10	3	106	0.9	0.008
2023-11	3	111	1.1	0.010
2023-12	3	115	1.3	0.011

(Continue sur la page suivante...)

TABLE 7 – Estimation des ratios correspondants aux revenus moyens pour la période allant de janvier 2023 à décembre 2023. (suite)

Période	Etoile	Estimateur	Ecart-type	cv
2023-01 2023-12	3	112	0.8	0.007
2023-01	4	162	2.7	0.017
2023-02	4	166	2.9	0.017
2023-03	4	164	2.1	0.013
2023-04	4	138	1.8	0.013
2023-05	4	144	1.8	0.013
2023-06	4	155	2.0	0.013
2023-07	4	140	1.7	0.012
2023-08	4	143	1.6	0.011
2023-09	4	150	1.6	0.011
2023-10	4	141	1.6	0.012
2023-11	4	144	1.5	0.010
2023-12	4	155	1.7	0.011
2023-01 2023-12	4	150	1.3	0.009
2023-01	5	424	12.9	0.030
2023-02	5	434	16.7	0.038
2023-03	5	384	7.7	0.020
2023-04	5	301	6.8	0.023
2023-05	5	367	8.8	0.024
2023-06	5	372	8.1	0.022
2023-07	5	364	7.4	0.020
2023-08	5	345	7.8	0.023
2023-09	5	325	5.5	0.017
2023-10	5	310	5.7	0.018
2023-11	5	317	6.7	0.021
2023-12	5	451	11.8	0.026
2023-01 2023-12	5	369	6.1	0.017

7 Limitations

Il est important de souligner que les CV sont assez petits, ceci indique une relativement bonne précision des estimations. Ceci est principalement dû au relativement bon taux de réponse, à la puissance de nos modèles de réponse et au calage performant. En outre, l'estimation de la covariance peut fluctuer et ainsi impacter l'estimation de variance de manière non-négligeable. Ces résultats sont donc encore à considérer comme expérimentaux.

Références

- CHEVALIER, C. & QUALITÉ, L. (2021). Rapport de méthodes : Enquête suisse sur la structure des salaires 2018, Révision de la pondération. Rapport de méthodes, Office Fédéral de la statistique OFS.
- DEVILLE, J.-C. (1999). Estimation de variance pour des statistiques et des estimateurs complexes : techniques de résidus et de linéarisation. *Techniques d'enquête* **25**, 219–230.
- DEVILLE, J.-C. & SÄRNDAL, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association* **87**, 376–382.
- KIM, J. K. & KIM, J. J. (2007). Nonresponse weighting adjustment using estimated response probability. *Canadian Journal of Statistics* **35**, 501–514.
- LE GUENNEC, J. & SAUTORY, O. (2002). CALMAR2 : une nouvelle version de la macro CALMAR de redressement d'échantillon par calage. In *Actes des Journées de Méthodologie, Insee, Paris*. Paris.
- TILLÉ, Y. (2020). *Théorie des sondages : Échantillonnage et estimation en populations finies*. Dunod, 2nd ed.

A Annexe - tel reminder et response channel

Analyse de l'impact des variables `tel_reminder` et `response_channel` pour l'utilisation dans le modèle de non-réponse.

```
proc glm data=output.HESTA_data_aug_periode(where=(rep2 = 1));
  class tel_reminder;
  model income = tel_reminder;
  ID OID;
  by instance_name;
run;
```

TABLE 8 – Résumé de la regression des revenus par l'indicateur de rappel.

INSTANCE_NAME	Dependent	RSquare	CV	RootMSE	DepMean
2022-08	INCOME	0	213	310 405	145 621
2022-09	INCOME	0	210	266 293	126 865
2022-10	INCOME	0	219	216 212	98 623
2022-11	INCOME	0	284	171 905	60 498
2022-12	INCOME	0	310	341 528	110 016
2023-01	INCOME	0	285	309 656	108 537
2023-02	INCOME	0	297	360 577	121 450
2023-03	INCOME	0	241	279 511	116 040
2023-04	INCOME	0	219	193 509	88 459
2023-05	INCOME	0	253	271 143	107 327
2023-06	INCOME	0	218	320 819	146 918

```
proc glm data=output.HESTA_data_aug_periode(where=(rep2 = 1));
  class response_channel;
  model income = response_channel;
  ID OID;
  by instance_name;
run;
```

TABLE 9 – Résumé de la regression des revenus par l'indicateur du canal de réponse.

INSTANCE_NAME	Dependent	RSquare	CV	RootMSE	DepMean
2022-08	INCOME	0.1	205	298 654	145 621
2022-09	INCOME	0.1	201	254 923	126 865
2022-10	INCOME	0.1	210	206 750	98 623
2022-11	INCOME	0.1	274	165 713	60 498
2022-12	INCOME	0.1	301	330 899	110 016
2023-01	INCOME	0.1	276	299 491	108 537
2023-02	INCOME	0.1	288	349 288	121 450
2023-03	INCOME	0.1	233	270 503	116 040
2023-04	INCOME	0.1	210	185 761	88 459
2023-05	INCOME	0.1	245	263 152	107 327
2023-06	INCOME	0.1	212	310 781	146 918

B Annexe - Évaluation des estimateurs

Le traitement de non-réponse alternatif intégrant le nombre de nuitées au deuxième niveau poserait quelques difficultés au niveau du nombre d'effectifs dans certains groupes homogènes. La table 10 montre l'incorporation des nuitées dans les groupes homogènes sous forme de variable catégorielle. Les figures 8 et 9 montrent que l'impact d'une telle adaptation est faible. Le nombre de nuitées n'est alors pas intégré au traitement de non-réponse.

TABLE 10 – Groupes homogènes de réponse avec incorporation des nuitées pour modéliser le mécanisme de non-réponse au revenu moyen par nuitée (deuxième niveau). Période 2023-01 à 2023-12.

Jours ouverts > 0	Étoiles	Nombre de lits	Zone touristique	Nombre de nuitées	min
0	-	-	-	-	212
1	5	-	-	-	81
1	4	1-70	1-2	-	68
1	4	71-150	1-2	-	83
1	4	1-70	3-4	-	46
1	4	71-150	3-4	-	80
1	4	151+	-	-	140
1	3	1-70	1-2	1	52
1	3	71-150	1-2	1	47
1	3	1-70	1-2	2	334
1	3	71-150	1-2	2	54
1	3	1-70	3-4	1	6
1	3	71-150	3-4	1	11
1	3	1-70	3-4	2	127
1	3	71-150	3-4	2	39
1	3	151+	-	-	62
1	100-0-1-2	1-70	1-2	1	585
1	100-0-1-2	71-150	1-2	1	50
1	100-0-1-2	1-70	1-2	2	541
1	100-0-1-2	71-150	1-2	2	25
1	100-0-1-2	1-70	3-4	1	87
1	100-0-1-2	71-150	3-4	1	14
1	100-0-1-2	1-70	3-4	2	190
1	100-0-1-2	71-150	3-4	2	22
1	100-0-1-2	151+	-	-	83

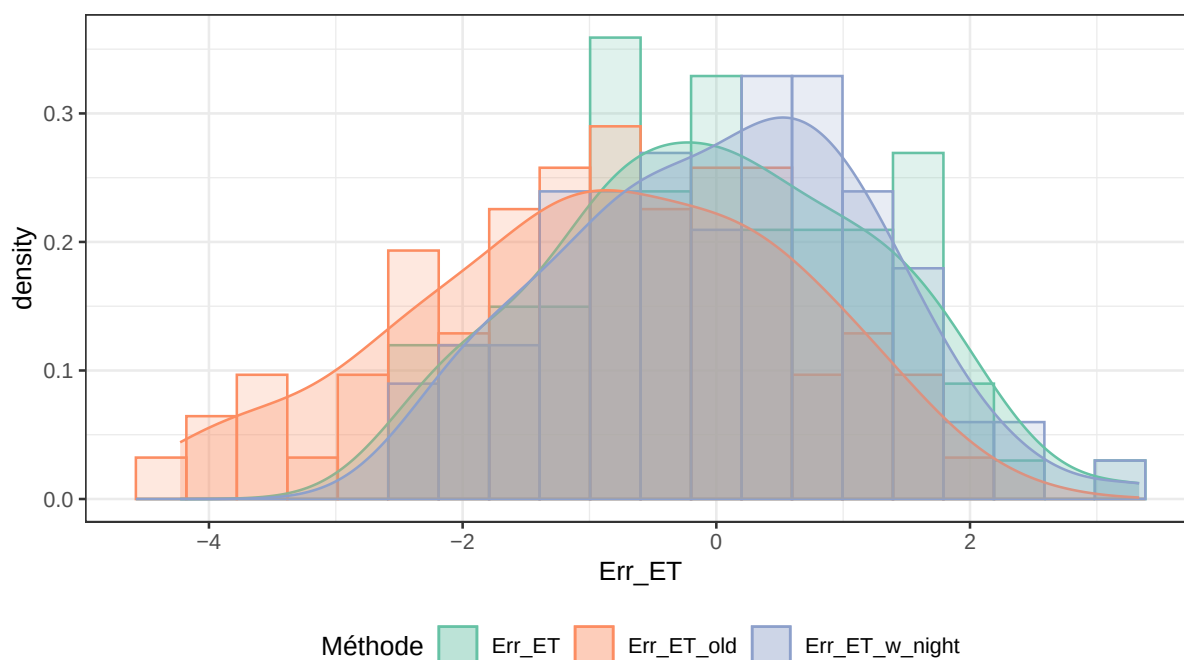


FIGURE 8 – Comparaison des histogrammes de l'écart-type des biais relatifs sur l'estimation des nuitées par rapport aux anciens groupes homogènes. Période 2022-05 à 2023-06.

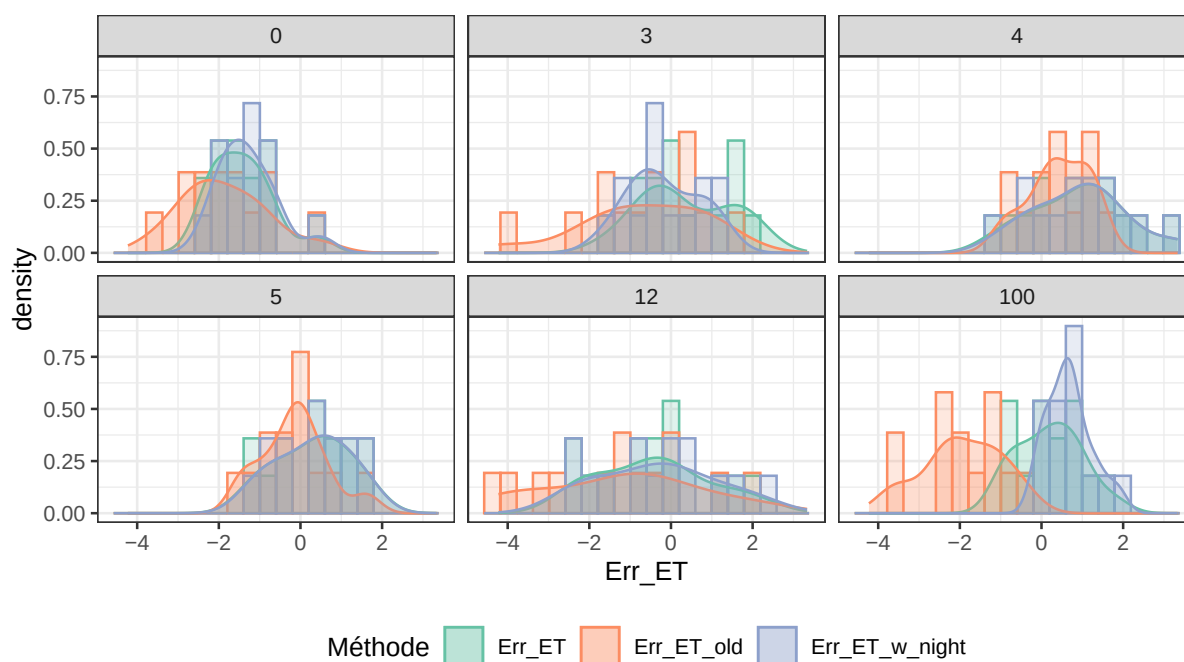


FIGURE 9 – Figure 8 détaillée par nombre d'étoiles.

C Annexe - Évaluation du calage

TABLE 11 – Table de vérification du nombre d’observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d’observations dans la strate, N le nombre d’observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux.

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06					16	3 211	4 505.0	4 505	287 519	287 519
2023-06				0-12	17	192	259.0	259	14 496	14 496
2023-06				0-34	17	43	51.0	51	5 612	5 612
2023-06				100-12	17	1 384	2 001.0	2 001	63 956	63 956
2023-06				100-34	17	230	330.0	330	17 393	17 393
2023-06				12-12	17	51	92.0	92	4 921	4 921
2023-06				12-34	17	25	49.0	49	6 988	6 988
2023-06				3-12	17	572	819.0	819	51 758	51 758
2023-06				3-34	17	195	264.0	264	25 514	25 514
2023-06				4-12	17	241	317.0	317	38 969	38 969
2023-06				4-34	17	179	212.0	212	37 566	37 566
2023-06				5-12	17	57	61.0	61	10 302	10 302
2023-06				5-34	17	42	50.0	50	10 044	10 044
2023-06		0			18	2 869	4 087.0	4 087	177 068	177 464
2023-06		1			18	342	418.0	418	110 451	110 055
2023-06		0		0-12	19	174	239.0	239	10 520	10 597
2023-06		0		0-34	19	34	41.7	41	2 210	2 174
2023-06		0		100-12	19	1 364	1 973.6	1 975	57 302	57 816
2023-06		0		100-34	19	213	308.5	309	11 532	11 807
2023-06		0		12-12	19	48	87.2	88	3 558	3 501
2023-06		0		12-34	19	13	27.4	28	1 656	1 558
2023-06		0		3-12	19	534	778.3	779	40 518	40 828
2023-06		0		3-34	19	163	226.0	224	15 461	15 053
2023-06		0		4-12	19	175	234.9	233	18 789	18 585
2023-06		0		4-34	19	102	115.9	116	10 437	10 471
2023-06		0		5-12	19	33	34.7	34	3 225	3 202
2023-06		0		5-34	19	16	19.7	21	1 860	1 872
2023-06		1		0-12	19	18	20.0	20	3 976	3 899
2023-06		1		0-34	19	9	9.3	10	3 402	3 438
2023-06		1		100-12	19	20	27.4	26	6 654	6 140
2023-06		1		100-34	19	17	21.5	21	5 861	5 586
2023-06		1		12-12	19	3	4.8	4	1 363	1 420
2023-06		1		12-34	19	12	21.6	21	5 332	5 430
2023-06		1		3-12	19	38	40.7	40	11 240	10 930
2023-06		1		3-34	19	32	38.0	40	10 053	10 461
2023-06		1		4-12	19	66	82.1	84	20 180	20 384
2023-06		1		4-34	19	77	96.1	96	27 129	27 095
2023-06		1		5-12	19	24	26.3	27	7 077	7 100
2023-06		1		5-34	19	26	30.3	29	8 184	8 172
2023-06	0				20	235	310.0	310	20 108	20 108
2023-06	3				20	767	1 083.0	1 083	77 272	77 272
2023-06	4				20	420	529.0	529	76 535	76 535
2023-06	5				20	99	111.0	111	20 346	20 346
2023-06	12				20	76	141.0	141	11 909	11 909
2023-06	100				20	1 614	2 331.0	2 331	81 349	81 349
2023-06	0			0-12	21	192	259.0	259	14 496	14 496
2023-06	0			0-34	21	43	51.0	51	5 612	5 612
2023-06	3			3-12	21	572	819.0	819	51 758	51 758
2023-06	3			3-34	21	195	264.0	264	25 514	25 514
2023-06	4			4-12	21	241	317.0	317	38 969	38 969
2023-06	4			4-34	21	179	212.0	212	37 566	37 566
2023-06	5			5-12	21	57	61.0	61	10 302	10 302

(Continue sur la page suivante...)

TABLE 11 – Table de vérification du nombre d'observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d'observations dans la strate, N le nombre d'observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux. (suite)

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06		5		5-34	21	42	50.0	50	10 044	10 044
2023-06		12		12-12	21	51	92.0	92	4 921	4 921
2023-06		12		12-34	21	25	49.0	49	6 988	6 988
2023-06		100		100-12	21	1 384	2 001.0	2 001	63 956	63 956
2023-06		100		100-34	21	230	330.0	330	17 393	17 393
2023-06		0	0		22	208	280.7	280	12 730	12 771
2023-06		0	1		22	27	29.3	30	7 378	7 337
2023-06		3	0		22	697	1 004.3	1 003	55 979	55 881
2023-06		3	1		22	70	78.7	80	21 293	21 391
2023-06		4	0		22	277	350.9	349	29 226	29 056
2023-06		4	1		22	143	178.1	180	47 309	47 479
2023-06		5	0		22	49	54.4	55	5 085	5 074
2023-06		5	1		22	50	56.6	56	15 261	15 272
2023-06		12	0		22	61	114.6	116	5 214	5 059
2023-06		12	1		22	15	26.4	25	6 695	6 850
2023-06		100	0		22	1 577	2 282.1	2 284	68 835	69 623
2023-06		100	1		22	37	48.9	47	12 514	11 726
2023-06		0	0	0-12	23	174	239.0	239	10 520	10 597
2023-06		0	0	0-34	23	34	41.7	41	2 210	2 174
2023-06		0	1	0-12	23	18	20.0	20	3 976	3 899
2023-06		0	1	0-34	23	9	9.3	10	3 402	3 438
2023-06		3	0	3-12	23	534	778.3	779	40 518	40 828
2023-06		3	0	3-34	23	163	226.0	224	15 461	15 053
2023-06		3	1	3-12	23	38	40.7	40	11 240	10 930
2023-06		3	1	3-34	23	32	38.0	40	10 053	10 461
2023-06		4	0	4-12	23	175	234.9	233	18 789	18 585
2023-06		4	0	4-34	23	102	115.9	116	10 437	10 471
2023-06		4	1	4-12	23	66	82.1	84	20 180	20 384
2023-06		4	1	4-34	23	77	96.1	96	27 129	27 095
2023-06		5	0	5-12	23	33	34.7	34	3 225	3 202
2023-06		5	0	5-34	23	16	19.7	21	1 860	1 872
2023-06		5	1	5-12	23	24	26.3	27	7 077	7 100
2023-06		5	1	5-34	23	26	30.3	29	8 184	8 172
2023-06		12	0	12-12	23	48	87.2	88	3 558	3 501
2023-06		12	0	12-34	23	13	27.4	28	1 656	1 558
2023-06		12	1	12-12	23	3	4.8	4	1 363	1 420
2023-06		12	1	12-34	23	12	21.6	21	5 332	5 430
2023-06		100	0	100-12	23	1 364	1 973.6	1 975	57 302	57 816
2023-06		100	0	100-34	23	213	308.5	309	11 532	11 807
2023-06		100	1	100-12	23	20	27.4	26	6 654	6 140
2023-06		100	1	100-34	23	17	21.5	21	5 861	5 586
2023-06	1				24	1 641	2 364.0	2 364	136 435	136 435
2023-06	2				24	856	1 185.0	1 185	47 967	47 967
2023-06	3				24	377	529.0	529	38 095	38 095
2023-06	4				24	337	427.0	427	65 022	65 022
2023-06	1			0-12	25	161	218.9	218	12 953	12 764
2023-06	1			100-12	25	798	1 178.4	1 177	42 608	42 505
2023-06	1			12-12	25	37	67.7	70	3 053	3 421
2023-06	1			3-12	25	411	598.9	596	38 338	38 441
2023-06	1			4-12	25	181	243.2	246	30 087	29 878
2023-06	1			5-12	25	53	56.9	57	9 396	9 426
2023-06	2			0-12	25	31	40.1	41	1 543	1 732
2023-06	2			100-12	25	586	822.6	824	21 348	21 451

(Continue sur la page suivante...)

TABLE 11 – Table de vérification du nombre d'observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d'observations dans la strate, N le nombre d'observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux. (suite)

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06	2			12-12	25	14	24.3	22	1 868	1 500
2023-06	2			3-12	25	161	220.1	223	13 420	13 317
2023-06	2			4-12	25	60	73.8	71	8 882	9 091
2023-06	2			5-12	25	4	4.1	4	906	876
2023-06	3			0-34	25	22	28.6	29	1 628	1 655
2023-06	3			100-34	25	139	212.9	215	8 248	8 348
2023-06	3			12-34	25	12	24.3	24	2 583	2 277
2023-06	3			3-34	25	105	149.6	148	10 686	10 570
2023-06	3			4-34	25	84	95.9	95	12 036	12 336
2023-06	3			5-34	25	15	17.7	18	2 913	2 909
2023-06	4			0-34	25	21	22.4	22	3 984	3 957
2023-06	4			100-34	25	91	117.1	115	9 145	9 045
2023-06	4			12-34	25	13	24.7	25	4 405	4 711
2023-06	4			3-34	25	90	114.4	116	14 828	14 944
2023-06	4			4-34	25	95	116.1	117	25 530	25 230
2023-06	4			5-34	25	27	32.3	32	7 131	7 135
2023-06	1	0			26	1 511	2 207.0	2 209	97 492	98 387
2023-06	1	1			26	130	157.0	155	38 943	38 048
2023-06	2	0			26	817	1 140.8	1 139	36 421	36 142
2023-06	2	1			26	39	44.2	46	11 546	11 825
2023-06	3	0			26	322	464.9	464	24 503	24 122
2023-06	3	1			26	55	64.1	65	13 592	13 973
2023-06	4	0			26	219	274.3	275	18 653	18 813
2023-06	4	1			26	118	152.7	152	46 369	46 209
2023-06	1	0		0-12	27	144	199.9	199	9 134	9 022
2023-06	1	0		100-12	27	783	1 157.1	1 157	37 192	37 604
2023-06	1	0		12-12	27	37	67.7	69	3 053	2 851
2023-06	1	0		3-12	27	384	569.5	568	30 328	30 841
2023-06	1	0		4-12	27	132	180.2	184	14 778	15 081
2023-06	1	0		5-12	27	31	32.6	32	3 006	2 988
2023-06	1	1		0-12	27	17	19.0	19	3 819	3 742
2023-06	1	1		100-12	27	15	21.4	20	5 416	4 901
2023-06	1	1		3-12	27	27	29.4	28	8 010	7 600
2023-06	1	1		4-12	27	49	63.0	62	15 309	14 797
2023-06	1	1		5-12	27	22	24.2	25	6 390	6 438
2023-06	2	0		0-12	27	30	39.1	40	1 386	1 575
2023-06	2	0		100-12	27	581	816.6	818	20 110	20 212
2023-06	2	0		12-12	27	11	19.5	19	504	650
2023-06	2	0		3-12	27	150	208.8	211	10 190	9 987
2023-06	2	0		4-12	27	43	54.7	49	4 011	3 504
2023-06	2	0		5-12	27	2	2.0	2	219	214
2023-06	2	1		0-12	27	1	1.0	1	157	157
2023-06	2	1		100-12	27	5	6.0	6	1 238	1 239
2023-06	2	1		12-12	27	3	4.8	3	1 363	850
2023-06	2	1		3-12	27	11	11.3	12	3 230	3 330
2023-06	2	1		4-12	27	17	19.1	22	4 871	5 587
2023-06	2	1		5-12	27	2	2.1	2	687	662
2023-06	3	0		0-34	27	20	26.6	27	1 248	1 275
2023-06	3	0		100-34	27	132	204.8	207	6 753	6 883
2023-06	3	0		12-34	27	7	16.0	16	831	713
2023-06	3	0		3-34	27	97	140.8	139	8 920	8 734
2023-06	3	0		4-34	27	59	68.2	66	5 806	5 637
2023-06	3	0		5-34	27	7	8.7	9	944	880

(Continue sur la page suivante...)

TABLE 11 – Table de vérification du nombre d'observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d'observations dans la strate, N le nombre d'observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux. (suite)

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06	3		1	0-34	27	2	2.0	2	380	380
2023-06	3		1	100-34	27	7	8.1	8	1 495	1 465
2023-06	3		1	12-34	27	5	8.4	8	1 753	1 564
2023-06	3		1	3-34	27	8	8.8	9	1 766	1 836
2023-06	3		1	4-34	27	25	27.7	29	6 230	6 699
2023-06	3		1	5-34	27	8	9.1	9	1 970	2 029
2023-06	4		0	0-34	27	14	15.1	14	961	899
2023-06	4		0	100-34	27	81	103.6	102	4 779	4 924
2023-06	4		0	12-34	27	6	11.5	12	826	845
2023-06	4		0	3-34	27	66	85.3	85	6 540	6 319
2023-06	4		0	4-34	27	43	47.8	50	4 630	4 834
2023-06	4		0	5-34	27	9	11.0	12	916	992
2023-06	4		1	0-34	27	7	7.3	8	3 022	3 058
2023-06	4		1	100-34	27	10	13.4	13	4 366	4 121
2023-06	4		1	12-34	27	7	13.2	13	3 579	3 866
2023-06	4		1	3-34	27	24	29.2	31	8 287	8 625
2023-06	4		1	4-34	27	52	68.3	67	20 900	20 396
2023-06	4		1	5-34	27	18	21.3	20	6 215	6 143
2023-06	1	0			28	161	218.9	218	12 953	12 764
2023-06	1	3			28	411	598.9	596	38 338	38 441
2023-06	1	4			28	181	243.2	246	30 087	29 878
2023-06	1	5			28	53	56.9	57	9 396	9 426
2023-06	1	12			28	37	67.7	70	3 053	3 421
2023-06	1	100			28	798	1 178.4	1 177	42 608	42 505
2023-06	2	0			28	31	40.1	41	1 543	1 732
2023-06	2	3			28	161	220.1	223	13 420	13 317
2023-06	2	4			28	60	73.8	71	8 882	9 091
2023-06	2	5			28	4	4.1	4	906	876
2023-06	2	12			28	14	24.3	22	1 868	1 500
2023-06	2	100			28	586	822.6	824	21 348	21 451
2023-06	3	0			28	22	28.6	29	1 628	1 655
2023-06	3	3			28	105	149.6	148	10 686	10 570
2023-06	3	4			28	84	95.9	95	12 036	12 336
2023-06	3	5			28	15	17.7	18	2 913	2 909
2023-06	3	12			28	12	24.3	24	2 583	2 277
2023-06	3	100			28	139	212.9	215	8 248	8 348
2023-06	4	0			28	21	22.4	22	3 984	3 957
2023-06	4	3			28	90	114.4	116	14 828	14 944
2023-06	4	4			28	95	116.1	117	25 530	25 230
2023-06	4	5			28	27	32.3	32	7 131	7 135
2023-06	4	12			28	13	24.7	25	4 405	4 711
2023-06	4	100			28	91	117.1	115	9 145	9 045
2023-06	1	0		0-12	29	161	218.9	218	12 953	12 764
2023-06	1	3		3-12	29	411	598.9	596	38 338	38 441
2023-06	1	4		4-12	29	181	243.2	246	30 087	29 878
2023-06	1	5		5-12	29	53	56.9	57	9 396	9 426
2023-06	1	12		12-12	29	37	67.7	70	3 053	3 421
2023-06	1	100		100-12	29	798	1 178.4	1 177	42 608	42 505
2023-06	2	0		0-12	29	31	40.1	41	1 543	1 732
2023-06	2	3		3-12	29	161	220.1	223	13 420	13 317
2023-06	2	4		4-12	29	60	73.8	71	8 882	9 091
2023-06	2	5		5-12	29	4	4.1	4	906	876
2023-06	2	12		12-12	29	14	24.3	22	1 868	1 500

(Continue sur la page suivante...)

TABLE 11 – Table de vérification du nombre d’observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d’observations dans la strate, N le nombre d’observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux. (suite)

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06	2	100		100-12	29	586	822.6	824	21 348	21 451
2023-06	3	0		0-34	29	22	28.6	29	1 628	1 655
2023-06	3	3		3-34	29	105	149.6	148	10 686	10 570
2023-06	3	4		4-34	29	84	95.9	95	12 036	12 336
2023-06	3	5		5-34	29	15	17.7	18	2 913	2 909
2023-06	3	12		12-34	29	12	24.3	24	2 583	2 277
2023-06	3	100		100-34	29	139	212.9	215	8 248	8 348
2023-06	4	0		0-34	29	21	22.4	22	3 984	3 957
2023-06	4	3		3-34	29	90	114.4	116	14 828	14 944
2023-06	4	4		4-34	29	95	116.1	117	25 530	25 230
2023-06	4	5		5-34	29	27	32.3	32	7 131	7 135
2023-06	4	12		12-34	29	13	24.7	25	4 405	4 711
2023-06	4	100		100-34	29	91	117.1	115	9 145	9 045
2023-06	1	0	0		30	144	199.9	199	9 134	9 022
2023-06	1	0	1		30	17	19.0	19	3 819	3 742
2023-06	1	3	0		30	384	569.5	568	30 328	30 841
2023-06	1	3	1		30	27	29.4	28	8 010	7 600
2023-06	1	4	0		30	132	180.2	184	14 778	15 081
2023-06	1	4	1		30	49	63.0	62	15 309	14 797
2023-06	1	5	0		30	31	32.6	32	3 006	2 988
2023-06	1	5	1		30	22	24.2	25	6 390	6 438
2023-06	1	12	0		30	37	67.7	69	3 053	2 851
2023-06	1	100	0		30	783	1 157.1	1 157	37 192	37 604
2023-06	1	100	1		30	15	21.4	20	5 416	4 901
2023-06	2	0	0		30	30	39.1	40	1 386	1 575
2023-06	2	0	1		30	1	1.0	1	157	157
2023-06	2	3	0		30	150	208.8	211	10 190	9 987
2023-06	2	3	1		30	11	11.3	12	3 230	3 330
2023-06	2	4	0		30	43	54.7	49	4 011	3 504
2023-06	2	4	1		30	17	19.1	22	4 871	5 587
2023-06	2	5	0		30	2	2.0	2	219	214
2023-06	2	5	1		30	2	2.1	2	687	662
2023-06	2	12	0		30	11	19.5	19	504	650
2023-06	2	12	1		30	3	4.8	3	1 363	850
2023-06	2	100	0		30	581	816.6	818	20 110	20 212
2023-06	2	100	1		30	5	6.0	6	1 238	1 239
2023-06	3	0	0		30	20	26.6	27	1 248	1 275
2023-06	3	0	1		30	2	2.0	2	380	380
2023-06	3	3	0		30	97	140.8	139	8 920	8 734
2023-06	3	3	1		30	8	8.8	9	1 766	1 836
2023-06	3	4	0		30	59	68.2	66	5 806	5 637
2023-06	3	4	1		30	25	27.7	29	6 230	6 699
2023-06	3	5	0		30	7	8.7	9	944	880
2023-06	3	5	1		30	8	9.1	9	1 970	2 029
2023-06	3	12	0		30	7	16.0	16	831	713
2023-06	3	12	1		30	5	8.4	8	1 753	1 564
2023-06	3	100	0		30	132	204.8	207	6 753	6 883
2023-06	3	100	1		30	7	8.1	8	1 495	1 465
2023-06	4	0	0		30	14	15.1	14	961	899
2023-06	4	0	1		30	7	7.3	8	3 022	3 058
2023-06	4	3	0		30	66	85.3	85	6 540	6 319
2023-06	4	3	1		30	24	29.2	31	8 287	8 625
2023-06	4	4	0		30	43	47.8	50	4 630	4 834

(Continue sur la page suivante...)

TABLE 11 – Table de vérification du nombre d'observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d'observations dans la strate, N le nombre d'observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux. (suite)

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06	4	4	1		30	52	68.3	67	20 900	20 396
2023-06	4	5	0		30	9	11.0	12	916	992
2023-06	4	5	1		30	18	21.3	20	6 215	6 143
2023-06	4	12	0		30	6	11.5	12	826	845
2023-06	4	12	1		30	7	13.2	13	3 579	3 866
2023-06	4	100	0		30	81	103.6	102	4 779	4 924
2023-06	4	100	1		30	10	13.4	13	4 366	4 121
2023-06	1	0	0	0-12	31	144	199.9	199	9 134	9 022
2023-06	1	0	1	0-12	31	17	19.0	19	3 819	3 742
2023-06	1	3	0	3-12	31	384	569.5	568	30 328	30 841
2023-06	1	3	1	3-12	31	27	29.4	28	8 010	7 600
2023-06	1	4	0	4-12	31	132	180.2	184	14 778	15 081
2023-06	1	4	1	4-12	31	49	63.0	62	15 309	14 797
2023-06	1	5	0	5-12	31	31	32.6	32	3 006	2 988
2023-06	1	5	1	5-12	31	22	24.2	25	6 390	6 438
2023-06	1	12	0	12-12	31	37	67.7	69	3 053	2 851
2023-06	1	100	0	100-12	31	783	1 157.1	1 157	37 192	37 604
2023-06	1	100	1	100-12	31	15	21.4	20	5 416	4 901
2023-06	2	0	0	0-12	31	30	39.1	40	1 386	1 575
2023-06	2	0	1	0-12	31	1	1.0	1	157	157
2023-06	2	3	0	3-12	31	150	208.8	211	10 190	9 987
2023-06	2	3	1	3-12	31	11	11.3	12	3 230	3 330
2023-06	2	4	0	4-12	31	43	54.7	49	4 011	3 504
2023-06	2	4	1	4-12	31	17	19.1	22	4 871	5 587
2023-06	2	5	0	5-12	31	2	2.0	2	219	214
2023-06	2	5	1	5-12	31	2	2.1	2	687	662
2023-06	2	12	0	12-12	31	11	19.5	19	504	650
2023-06	2	12	1	12-12	31	3	4.8	3	1 363	850
2023-06	2	100	0	100-12	31	581	816.6	818	20 110	20 212
2023-06	2	100	1	100-12	31	5	6.0	6	1 238	1 239
2023-06	3	0	0	0-34	31	20	26.6	27	1 248	1 275
2023-06	3	0	1	0-34	31	2	2.0	2	380	380
2023-06	3	3	0	3-34	31	97	140.8	139	8 920	8 734
2023-06	3	3	1	3-34	31	8	8.8	9	1 766	1 836
2023-06	3	4	0	4-34	31	59	68.2	66	5 806	5 637
2023-06	3	4	1	4-34	31	25	27.7	29	6 230	6 699
2023-06	3	5	0	5-34	31	7	8.7	9	944	880
2023-06	3	5	1	5-34	31	8	9.1	9	1 970	2 029
2023-06	3	12	0	12-34	31	7	16.0	16	831	713
2023-06	3	12	1	12-34	31	5	8.4	8	1 753	1 564
2023-06	3	100	0	100-34	31	132	204.8	207	6 753	6 883
2023-06	3	100	1	100-34	31	7	8.1	8	1 495	1 465
2023-06	4	0	0	0-34	31	14	15.1	14	961	899
2023-06	4	0	1	0-34	31	7	7.3	8	3 022	3 058
2023-06	4	3	0	3-34	31	66	85.3	85	6 540	6 319
2023-06	4	3	1	3-34	31	24	29.2	31	8 287	8 625
2023-06	4	4	0	4-34	31	43	47.8	50	4 630	4 834
2023-06	4	4	1	4-34	31	52	68.3	67	20 900	20 396
2023-06	4	5	0	5-34	31	9	11.0	12	916	992
2023-06	4	5	1	5-34	31	18	21.3	20	6 215	6 143
2023-06	4	12	0	12-34	31	6	11.5	12	826	845
2023-06	4	12	1	12-34	31	7	13.2	13	3 579	3 866
2023-06	4	100	0	100-34	31	81	103.6	102	4 779	4 924

(Continue sur la page suivante...)

TABLE 11 – Table de vérification du nombre d’observations au sein de certains sous-groupes calés illustré pour le mois de juin 2023. La variable n représente le nombre d’observations dans la strate, N le nombre d’observations dans le cadre, $\hat{N}$ la somme des poids dans la strate, x_j le nombre de lits dans le cadre, $\hat{x}_j$ la somme du nombre de lits multiplié par les poids de calage. Les lignes en gris correspondent aux variables de calage retenues, les estimations et les totaux du cadre sont donc égaux. (suite)

Mois	Zone	Etoile	Big bed	Etoile x zone	type	n	$\hat{N}$	N	$\hat{x}_j$	x_j
2023-06	4	100	1	100-34	31	10	13.4	13	4 366	4 121