

VARIABILITÉ D'ÉCHANTILLONNAGE

L'Enquête sur la population active, sur laquelle reposent les estimations du disque compact, recueille des renseignements auprès d'un échantillon de ménages. Des estimations différentes pourraient avoir été obtenues d'un recensement complet des ménages en utilisant des questionnaires, des intervieweurs, des méthodes de traitement, etc. d'une façon identique à ce qui est appliqué présentement dans l'Enquête sur la Population Active.

L'erreur d'échantillonnage, ou erreur-type, est une mesure permettant de quantifier l'écart entre une estimation fondée sur un échantillon et la valeur tirée du recensement. Elle se base sur la notion du prélèvement de plusieurs échantillons, bien que, dans une enquête, on prélève un seul échantillon et on recueille des renseignements se rapportant aux unités de cet échantillon. Le prélèvement d'un grand nombre d'échantillons à partir d'une population donnée, selon le même plan d'échantillonnage, produirait une estimation située à une erreur-type de la valeur censitaire dans environ 68 % des échantillons et à deux erreurs-type de la valeur censitaire dans environ 95% des échantillons.

On peut aussi exprimer la variabilité d'échantillonnage en fonction de l'estimation. Le coefficient de variation (CV) est une mesure de l'erreur d'échantillonnage définie en pourcentage de l'estimation. En fait, il s'agit d'une erreur-type relative. Il est également possible d'établir des probabilités en utilisant le CV. Par exemple, avec un CV de 7 %, on peut dire que dans 68 % des échantillons, la valeur censitaire se trouvera au maximum à plus ou moins 7 % (la valeur du CV) de l'estimation alors que dans 95 % des échantillons, elle se situera au maximum à plus ou moins 14 % (le double du CV) de l'estimation. En ce qui concerne les estimations contenues sur ce disque compact, le CV est utilisé afin de donner une idée de l'incertitude associée aux estimations.

INDICATEURS DE LA VARIABILITÉ D'ÉCHANTILLONNAGE

Il est préférable d'obtenir des CV peu élevés puisque de tels CV indiquent que la variabilité d'échantillonnage est faible par rapport à l'estimation. Le CV dépend de la valeur des estimations, de la taille de l'échantillon à partir duquel l'estimation est établie, de la répartition de l'échantillon et de l'utilisation d'estimations postcensitaires de la population dans la procédure d'estimation. La valeur des estimations est importante parce que le CV représente l'erreur d'échantillonnage exprimée en pourcentage de l'estimation. Plus la valeur de l'estimation est faible, plus le CV est élevé (toutes choses étant égales par ailleurs). Par exemple, lorsque le taux de chômage est élevé, le CV peut être faible. Si le taux de chômage baisse en raison de l'amélioration de la situation économique, le CV correspondant augmentera. Normalement, dans le cas d'estimations similaires, l'estimation fondée sur l'échantillon le plus important est associée au CV le plus bas, car l'erreur d'échantillonnage est plus petite.

Par ailleurs, les estimations se rapportant à des caractéristiques plus groupées donnent lieu à un CV plus élevé. Ainsi, les personnes employées dans les secteurs de la foresterie, de la pêche, de l'exploitation minière et de l'exploitation pétrolière et gazière au Canada sont plus regroupées sur le plan géographique que les femmes de 55 à 64 ans employées en Ontario. La variabilité d'échantillonnage correspondant à ce dernier groupe sera plus faible, même si la valeur des estimations est similaire.

Enfin, les estimations se rapportant à l'âge et au sexe sont généralement plus fiables que d'autres estimations analogues parce que, dans le cas des estimations de l'EPA, l'échantillon est calibré en fonction des estimations postcensitaires de la population selon diverses catégories d'âge et de sexe. Pour poursuivre l'exemple précédent, les personnes employées à temps partiel en Alberta seront associées à une variabilité d'échantillonnage plus forte que les hommes de 35 à 44 ans employés en Colombie-Britannique, même si la valeur des estimations est similaire.

Variabilité des estimations mensuelles

Pour déterminer le CV approximatif d'une estimation mensuelle, veuillez consulter le tableau A, lequel présente l'estimation en fonction de la région géographique et du CV. Les lignes correspondent à la région géographique à laquelle l'estimation est associée et les colonnes, au degré de précision exprimé en CV, selon la grandeur de l'estimation. Pour déterminer le CV d'une estimation X dans la région A, suivez la ligne de la région A jusqu'à l'estimation la plus proche de X sans la dépasser. Le titre de la colonne indiquera le CV approximatif. Par exemple, pour connaître l'erreur d'échantillonnage d'une estimation de 38 000 chômeurs à Terre-Neuve-et-Labrador en septembre 2002, on constate que l'estimation inférieure la plus proche, de 28 600, donne un CV de 5 %. L'estimation de 38 000 chômeurs à Terre-Neuve-et-Labrador aura donc un CV d'environ 5 %.

Tableau A: CV pour les estimations* mensuelles, Canada et provinces

Région géographique	Coefficient de variation								
	1,0 %	2,5 %	5,0 %	7,5 %	10,0 %	15,0 %	20,0 %	25,0 %	30,0 %
Canada	1 093,5	311,4	134,2	77,1	42,8	24,6	16,3	11,7	8,9
Terre-Neuve-et-Labrador	265,0	70,9	28,6	15,9	8,9	4,9	3,2	2,3	1,7
Île-du-Prince-Édouard	69,4	20,0	8,5	4,9	2,8	1,6	1,1	0,8	0,6
Nouvelle-Écosse	243,9	70,6	30,8	17,8	10,0	5,8	3,8	2,8	2,1
Nouveau-Brunswick	221,2	61,5	25,5	14,4	8,2	4,7	3,1	2,2	1,7
Québec	1 183,8	323,3	132,3	74,2	41,9	23,5	15,3	10,9	8,2
Ontario	1 033,3	294,1	126,7	72,7	40,4	23,2	15,3	11,0	8,4
Manitoba	214,5	64,4	29,3	17,3	9,6	5,7	3,8	2,8	2,1
Saskatchewan	219,1	61,7	26,4	15,1	8,3	4,8	3,1	2,3	1,7
Alberta	545,0	162,7	74,0	43,6	24,0	14,1	9,5	6,9	5,3
Colombie-Britannique	784,9	222,8	95,8	55,0	30,5	17,5	11,5	8,3	6,3

* Les estimations sont en milliers.

Le tableau A donne une idée approximative de la variabilité d'échantillonnage. Cette dernière est modélisée pour que le CV de l'estimation soit inférieur ou égal au CV du tableau dans environ 75 % des cas. Néanmoins, dans 25 % des cas, le CV de l'estimation sera plus élevé que celui indiqué au tableau.

La mesure de la variabilité qui apparaît au tableau A correspond à une moyenne établie à partir des données de l'échantillon de l'EPA de 1999, 2000, 2001 et de la majeure partie de 2002. Il faut bien se rappeler que ce CV est une approximation seulement. Pour une mesure plus précise de la variabilité, veuillez composez notre numéro sans frais 1 866 873-8788 ou communiquez avec nous par courriel à travail@statcan.ca.

Variabilité des estimations annuelles

Pour déterminer le CV approximatif d'une estimation d'une moyenne annuelle, veuillez consulter le tableau B, lequel présente l'estimation en fonction de- la région géographique et du CV. Les lignes correspondent à la région géographique à laquelle l'estimation est associée et les colonnes, au degré de précision exprimé en CV, selon la grandeur de l'estimation. Pour déterminer le CV d'une estimation X dans la région A, suivez la ligne de la région A jusqu'à l'estimation la plus proche de X sans la dépasser. Le titre de la colonne indiquera le CV approximatif. Par exemple, pour connaître l'erreur d'échantillonnage d'une estimation de 41 000 chômeurs à Terre-Neuve-et-Labrador, on constate que l'estimation inférieure la plus proche, de 19 900, donne un CV de 2,5 %. L'estimation de 41 000 chômeurs à Terre-Neuve-et-Labrador aura donc un CV d'environ 2,5 %.

Tableau B: CV pour les estimations* de moyennes annuelles, Canada et provinces

Région géographique	Coefficient de variation								
	1,0 %	2,5 %	5,0 %	7,5 %	10,0 %	15,0 %	20,0 %	25,0 %	30,0 %
Canada	384,5	112,6	51,1	30,0	16,1	9,4	6,3	4,6	3,5
Terre-Neuve-et-Labrador	71,4	19,9	9,0	5,2	2,6	1,5	1,0	0,7	0,5
Île-du-Prince-Édouard	17,0	5,3	2,6	1,6	0,9	0,5	0,4	0,3	0,2
Nouvelle-Écosse	64,6	20,2	10,0	6,1	3,2	1,9	1,3	1,0	0,8
Nouveau-Brunswick	57,4	16,6	7,8	4,6	2,3	1,4	0,9	0,7	0,5
Québec	297,0	88,3	41,7	24,9	12,9	7,7	5,2	3,7	2,9
Ontario	278,5	86,9	43,1	26,4	13,7	8,4	5,7	4,2	3,3
Manitoba	72,9	22,1	10,8	6,5	3,3	2,0	1,4	1,0	0,8
Saskatchewan	66,0	18,3	8,4	4,9	2,4	1,4	0,9	0,6	0,5
Alberta	179,6	53,6	25,9	15,6	7,9	4,7	3,2	2,3	1,8
Colombie-Britannique	209,3	63,0	30,6	18,5	9,4	5,7	3,8	2,8	2,1

* Les estimations sont en milliers.

Le tableau B donne une idée approximative de la variabilité d'échantillonnage. Cette dernière est modélisée pour que le CV de l'estimation soit inférieur ou égal au CV du tableau dans environ 75 % des cas. Néanmoins, dans 25 % des cas, le CV de l'estimation sera plus élevé que celui indiqué au tableau.

La mesure de la variabilité qui apparaît au tableau B correspond à une moyenne établie à partir des données l'échantillon de l'EPA de 1997, 1998, 1999, 2000 et de la majeure partie de 2001. Il faut bien se rappeler que ce CV est une approximation seulement. Pour une mesure plus précise de la variabilité, veuillez composez notre numéro sans frais 1 866 873-8788 ou communiquez avec nous par courriel à travail@statcan.ca.

Variabilité des taux

Pour les estimations exprimées sous forme de taux ou de pourcentage, la variabilité d'échantillonnage dépend de la variabilité du numérateur et du dénominateur du rapport. Les divers taux donnés sur le disque compact sont traités différemment, car certains dénominateurs sont des valeurs étalonnées pour lesquelles il n'y a pas de variabilité d'échantillonnage.

Taux de chômage

Dans un groupe de personnes, le taux de chômage est défini comme le rapport du nombre de chômeurs, soit X, sur le nombre total de personnes dans la population active, soit Y. Le groupe peut désigner une province ou une RMR et(ou) un groupe d'âge-sexe. Par exemple, en septembre 2002, on dénombrait environ 38 000 chômeurs à Terre-Neuve-et-Labrador, tandis que la population active de la province comptait 259 000 personnes, ce qui donnait un taux de chômage de 14,6 %.

Pour déterminer le CV du taux de chômage, on se sert de la formule suivante :

$$CV(X/Y) = \sqrt{[CV(X)]^2 - [CV(Y)]^2}$$

où CV(X) est la valeur du CV associée au nombre total de chômeurs d'un sous-groupe régional ou démographique et CV(Y), la valeur du CV associé au nombre total de personnes du même sous-groupe faisant partie de la population active. (Ne devrait-on pas plutôt dire: où CV(X) est la valeur du CV associée au nombre total de chômeurs... et CV(Y), la valeur du CV associée au nombre total de personnes...) Pour reprendre l'exemple de Terre-Neuve-et-Labrador, le CV du taux de chômage de 14,6 % serait égal à $\sqrt{5^2 - 1^2} = 4,9\%$; le CV associé aux estimations mensuelles de 38 000 et de 259 000 à Terre-Neuve-et-Labrador est respectivement de 5 % et de 1 %, selon le tableau A.

Taux d'activité et taux d'emploi

Le taux d'activité représente la population active exprimée en pourcentage de la population totale. Le taux d'emploi est le nombre total de personnes occupées divisé par la population totale. Le numérateur et le dénominateur de ces deux rapports se rapportent au même groupe régional et démographique.

Les estimations de la population de l'EPA du Canada, des provinces, des RMR et des groupes d'âge-sexe n'est pas sujette à la variabilité d'échantillonnage, car elles sont étalonnées selon des sources indépendantes. Par conséquent, le CV du taux d'activité et du taux d'emploi des groupes régionaux et démographiques correspond au CV du numérateur.

On appelle «domaine» les sous-groupes se trouvant à l'intérieur du Canada, des provinces et des groupes d'âge-sexe. Ainsi, les personnes occupées dans la branche de l'agriculture au Manitoba forment un domaine. Pour déterminer le CV des taux se rapportant à un domaine, on doit tenir compte de la variabilité du numérateur et du dénominateur, car le dénominateur n'est plus un total contrôlé. Il est donc sujet à la variabilité d'échantillonnage. On peut calculer le CV associé au taux d'activité et au taux d'emploi dans un domaine de la même façon que celle montrée précédemment pour le taux de chômage. Les totaux figurant au numérateur et au dénominateur d'un taux en particulier devraient se rapporter au même domaine ou au même sous-groupe.

Variabilité de l'estimation d'un changement

La différence entre des estimations provenant de deux périodes différentes donne une estimation du changement, laquelle subit aussi la variabilité d'échantillonnage. L'estimation d'un changement entre deux années ou entre deux mois repose sur deux échantillons qui peuvent avoir des ménages en commun. Ainsi, le CV du changement dépend à la fois du CV des estimations des deux périodes et du chevauchement entre les échantillons des deux périodes, ρ . La formule qui suit peut être utilisée pour calculer le CV de l'estimation du changement :

$$CV(Y_2 - Y_1) = \sqrt{1 - \rho} \frac{\sqrt{Y_1^2 CV(Y_1)^2 + Y_2^2 CV(Y_2)^2}}{(Y_2 - Y_1)}$$

où Y_1 et Y_2 correspondent à l'estimation respective des deux périodes; la valeur de ρ est 0,5 pour un changement entre deux mois consécutifs et zéro pour un changement couvrant toute autre période. Lorsqu'on compare les moyennes annuelles de deux années, les CV des estimations annuelles devraient être dérivées à partir du tableau B. Pour un changement entre deux mois, les estimations désaisonnalisées devraient être utilisées, conjointement avec les CV des estimations mensuelles dérivées à partir du tableau A. Veuillez prendre note que la formule ci-dessus produit des estimations approximatives de la variabilité de l'échantillonnage associées à l'estimation d'un changement.