

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. La différence entre les estimations provenant d'un échantillon et celles d'un recensement s'appelle l'erreur d'échantillonnage.

L'erreur d'échantillonnage ou l'erreur-type est une mesure fondée sur le principe de la sélection de plusieurs échantillons, malgré le fait que, dans le cadre d'une enquête, un seul échantillon est sélectionné et l'information recueillie auprès d'unités choisies au sein 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

Un faible CV est préférable, car il indique une petite variabilité de l'échantillonnage par rapport à l'estimation. Le CV dépend de la grandeur de l'estimation, de la taille de l'échantillon qui a servi à calculer cette dernière et de la distribution de l'échantillon. Entre deux estimations, celle reposant sur le plus grand échantillon se caractérisera sans doute par la plus petite erreur d'échantillonnage. Entre deux estimations issues d'un échantillon de même importance, celle qui se rapporte au paramètre géographiquement le plus regroupé présentera la plus grande variabilité. Par exemple, au Canada, les personnes qui travaillent dans la branche de l'exploitation forestière se caractériseront par une plus grande concentration géographique que les veuves de 45 ans et plus qui sont occupées. Le second groupe présentera donc une plus petite variabilité de l'échantillonnage, même si les estimations ont approximativement la même importance. En outre, les estimations liées à l'âge et au sexe sont souvent plus fiables que d'autres dérivant d'un échantillon de taille comparable car, dans l'EPA, l'échantillon est étalonné selon l'âge et le sexe d'après des sources indépendantes. Poursuivant l'exemple précédent, au Canada, les veuves de 45 ans et plus qui sont occupées auront une plus grande variabilité d'échantillonnage que les personnes en chômage âgées de 55 ans et plus, même si les estimations sont de taille semblable.

Variabilité des estimations mensuelles

Pour déterminer le CV approximatif d'une estimation mensuelles, veuillez consulter le tableau A, lequel présente l'estimation en fonction des paramètres géographiques et du CV. Les lignes correspondent à la région de l'estimation 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 34 000 chômeurs à Terre-Neuve-et-Labrador en septembre 2001, on constate que l'estimation inférieure la plus proche, de 29 300, donne un CV de 5 %. L'estimation de 34 000 chômeurs 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	1 089,6	308,1	132,1	75,6	41,9	24,0	15,8	11,3	8,6
Terre-Neuve-et-Labrador	280,1	73,6	29,3	16,2	9,0	4,9	3,2	2,2	1,7
Île-du-Prince-Édouard	87,2	23,5	9,3	5,2	3,0	1,7	1,1	0,8	0,6
Nouvelle-Écosse	245,6	70,7	30,7	17,7	9,9	5,7	3,8	2,7	2,1
Nouveau-Brunswick	263,2	69,8	27,3	15,1	8,7	4,8	3,1	2,2	1,7
Québec	1 182,6	324,2	133,0	74,8	42,3	23,8	15,5	11,1	8,4
Ontario	1 024,3	289,6	124,2	71,1	39,4	22,5	14,9	10,7	8,1
Manitoba	223,6	66,2	29,9	17,6	9,6	5,7	3,8	2,7	2,1
Saskatchewan	227,0	63,0	26,7	15,1	8,3	4,7	3,1	2,2	1,7
Alberta	535,6	159,4	72,3	42,6	23,4	13,8	9,2	6,7	5,1
Colombie-Britannique	804,3	224,1	95,2	54,1	29,8	16,9	11,1	7,9	6,0

* Les estimations sont en milliers.

Le tableau A donne une idée approximative de la variabilité de l'échantillonnage. Cette dernière est modélisée pour que le CV soit inférieur ou égal au CV du tableau dans environ 75 % des cas, pour une estimation donnée. Dans 25 % des cas, l'estimation sera néanmoins légèrement plus élevée que celle indiquée au tableau.

La mesure de la variabilité qui apparaît au tableau A correspond à une moyenne pour le plan d'échantillonnage et la taille de l'échantillon de 1997, de 1998, 1999, 2000 et de la majeure partie de 2001. Il faut bien se rappeler que le CV est approximatif. Il n'illustre que la variabilité de l'échantillonnage. Pour une mesure plus précise de la variabilité, veuillez communiquer avec la Sous-division de l'enquête sur la population active, Statistique Canada, Ottawa, KIA OT6.

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 des paramètres géographiques et du CV. Les lignes correspondent à la région de l'estimation 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 21 100, donne un CV de 2,5 %. L'estimation de 41 000 chômeurs 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	407,1	114,8	50,3	28,9	15,5	8,9	5,9	4,2	3,2
Terre-Neuve-et-Labrador	77,3	21,1	9,4	5,4	2,7	1,5	1,0	0,7	0,5
Île-du-Prince-Édouard	19,2	5,8	2,7	1,6	0,8	0,5	0,3	0,2	0,2
Nouvelle-Écosse	66,3	20,3	10,0	6,0	3,1	1,9	1,3	0,9	0,7
Nouveau-Brunswick	60,3	17,3	8,0	4,7	2,4	1,4	0,9	0,7	0,5
Québec	303,0	89,7	42,3	25,2	13,0	7,7	5,2	3,8	2,9
Ontario	277,7	84,9	41,7	25,3	13,0	7,9	5,4	3,9	3,0
Manitoba	76,5	22,8	11,1	6,6	3,4	2,0	1,4	1,0	0,8
Saskatchewan	71,0	19,1	8,7	5,0	2,4	1,4	0,9	0,6	0,5
Alberta	174,9	52,5	25,5	15,4	7,8	4,7	3,2	2,3	1,8
Colombie-Britannique	212,4	63,0	30,3	18,2	9,2	5,5	3,7	2,7	2,0

* Les estimations sont en milliers.

Le tableau B donne une idée approximative de la variabilité de l'échantillonnage. Cette dernière est modélisée pour que le CV soit inférieur ou égal au CV du tableau dans environ 75 % des cas, pour une estimation donnée. Dans 25 % des cas, l'estimation sera néanmoins légèrement plus élevée que celle indiquée au tableau.

La mesure de la variabilité qui apparaît au tableau B correspond à une moyenne pour le plan d'échantillonnage et la taille de l'échantillon de 1997, 1998, 1999, 2000 et de la majeure partie de 2001. Il faut bien se rappeler que le CV est approximatif. Il illustre que la variabilité de l'échantillonnage. Pour une mesure plus précise de la variabilité, veuillez communiquer avec la Sous-division de l'enquête sur la population active, Statistique Canada, Ottawa, KIA OT6.

Variabilité des taux

Pour les estimations exprimées sous forme de taux ou de pourcentage, la variabilité de l'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 qui échappent à la variabilité de l'échantillonnage.

Taux de chômage

Le taux de chômage correspond au rapport de x , représentant le nombre total de chômeurs dans un groupe, et de y , nombre total de membres du même groupe faisant partie de la population active. Le groupe peut désigner une province ou une RMR et (ou) un groupe d'âge-sexe. En septembre 2001, par exemple, on dénombrait environ 34 000 chômeurs à Terre-Neuve-et-Labrador, tandis que la population active de la province comptait 254 000 personnes, ce qui donne un taux de chômage de 12,6 %.

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 x représente le nombre total de chômeurs d'un sous-groupe régional ou démographique et y , le nombre total de personnes du même sous-groupe faisant partie de la population active. Pour reprendre l'exemple de Terre-Neuve-et-Labrador, le CV du taux de chômage de 12,6 % serait de $\sqrt{5^2 - 1^2} = 4,9\%$, le CV des estimations mensuelles de 34 000 et de 254 000 à Terre-Neuve-et-Labrador étant 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 désignent le même groupe régional et démographique.

La population du Canada, des provinces, des RMR et des groupes d'âge-sexe n'est pas sujette à la variabilité de l'échantillonnage, car elle est étalonnée 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 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 rapports d'un domaine, on doit tenir compte de la variabilité du numérateur et du dénominateur, car le dénominateur n'est plus un total contrôlé. Il subit donc la variabilité de l'échantillonnage. Pour le taux d'activité et le taux d'emploi des domaines, on peut calculer le CV de la même façon que pour le taux de chômage. Le total apparaissant au numérateur du taux pertinent et celui figurant au dénominateur 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 de l'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 (tableau B) devraient être utilisés. Pour un changement entre deux mois, les estimations désaisonnalisées devraient être utilisées, conjointement avec les CV des estimations mensuelles (tableau A). Veuillez prendre note que la formule ci-dessus produit des estimations approximatives.