

10

**RAPPORT SUR LA QUALITÉ
DES DONNÉES**

Introduction

Le coefficient de variation d'une statistique donnée est largement tributaire de la taille de l'échantillon global de l'enquête et de l'importance de la sous-population par rapport à la population canadienne totale. Il dépend également du taux de non-réponse et du plan de sondage adopté. Le rapport présente, à titre d'information générale, des tableaux affichant les tailles d'échantillon selon la province de même que les taux de réponse mensuels selon la province.

Ce chapitre précise la façon d'obtenir le coefficient de variation approximatif d'une statistique donnée à partir des tableaux des degrés approximatifs de variabilité de l'échantillonnage pour l'Enquête sur les voyages des Canadiens.

***Taille de
l'échantillon
de l'EVC***

Le tableau suivant indique le nombre de membres des ménages qui, dans les groupes de renouvellement échantillonnés de l'EPA, étaient admissibles à l'enquête supplémentaire, soit l'Enquête sur les voyages des Canadiens.

TABLEAU 6. *Taille des échantillons mensuels portant sur les voyages intraprovinciaux, selon la province, 1997*

	Jan.	Fév.	Mar.	Avr.	Mai	Juin
T.-N.	304	286	317	299	296	315
Î.-P.-É.	230	226	233	238	257	230
N.-É.	558	556	557	565	554	594
N.-B.	540	526	520	548	497	524
QC	1684	1658	1690	1679	1685	1726
Ont.	2462	2519	2605	2589	2564	2588
Man.	624	597	606	658	583	589
Sask.	530	561	535	519	541	552
Alb.	655	659	654	660	640	613
C.-B.	782	742	748	748	790	741
Canada	8369	8330	8465	8503	8407	8472

TABLEAU 6.
*(suite)***Taille des échantillons mensuels portant sur les voyages intraprovinciaux, selon la province, 1997**

	Juil.	Août	Sept.	Oct.	Nov.	Déc.
T.-N.	306	297	325	330	322	322
Î.-P.-É.	235	246	245	215	216	242
N.-É.	574	605	554	554	586	537
N.-B.	509	557	538	479	498	480
QC	1714	1820	1777	1725	1673	1574
Ont.	2635	2666	2601	2552	2452	2445
Man.	650	624	611	641	646	640
Sask.	538	554	563	641	641	678
Alb.	656	668	653	656	651	653
C.-B.	823	761	761	785	798	783
Canada	8640	8798	8628	8578	8483	8354

Nota: Ce sous-échantillon représente un groupe de renouvellement de l'échantillon de l'EPA

TABLEAU 6.1 *Taille des échantillons mensuels portant sur les voyages interprovinciaux, selon la province, 1997*

	Jan.	Fév.	Mar.	Avr.	Mai	Juin
T.-N.	940	901	936	892	935	915
Î.-P.-É.	682	685	700	703	735	701
N.-É.	1728	1648	1715	1699	1699	1744
N.-B.	1609	1598	1584	1616	1561	1618
QC	5222	5182	5126	5089	5074	5099
Ont.	7740	7767	7793	7792	7718	7801
Man.	1850	1868	1857	1885	1860	1854
Sask.	1641	1652	1628	1649	1635	1645
Alb.	1934	2002	1975	1982	1970	1952
C.-B.	2384	2295	2347	2259	2347	2267
Canada	25730	25598	25661	25566	25534	25596

TABLEAU 6.1
*(suite)***Taille des échantillons mensuels portant sur les voyages interprovinciaux, selon la province, 1997**

	Juil.	Août	Sept.	Oct.	Nov.	Déc.
T.-N.	921	913	927	931	962	936
Î.-P.-É.	727	721	740	692	692	705
N.-É.	1685	1788	1703	1768	1749	1700
N.-B.	1527	1635	1552	1572	1547	1522
QC	5117	5311	5259	5354	5203	4941
Ont.	7848	7981	7929	7913	7730	7731
Man.	1849	1896	1860	1868	1914	1899
Sask.	1610	1643	1657	1778	1752	1879
Alb.	1947	1976	1998	1970	1992	1997
C.-B.	2391	2284	2411	2334	2424	2361
Canada	25622	26148	26036	26180	25965	25671

Nota: Ce sous-échantillon représente trois groupes de renouvellement de l'échantillon de l'EPA

Taux de réponse de l'EVC

Le tableau suivant présente les taux de réponse de l'Enquête sur les voyages des Canadiens de 1997. Les taux de réponse indiqués correspondent à la proportion des personnes admissibles à l'Enquête sur les voyages des Canadiens qui ont fourni des renseignements. Ces taux de réponse ne sont pas cumulatifs, c'est-à-dire qu'ils ne tiennent pas compte des personnes qui auraient été admissibles à l'EVC mais qui n'ont pas pris part à l'EPA. En effet, on ne demande pas aux personnes qui ne participent pas à l'EPA si elles veulent répondre aux questions de l'EVC. On ne peut donc pas les considérer comme des non-répondants à l'EVC.

TABLEAU 7. *Taux de réponse mensuels selon la province, en pourcentage, 1997*

	Jan.	Fév.	Mar	Avr.	Mai	Juin
T.-N.	91.8	90.6	92.8	92.4	92.5	91.0
Î.-P.-É.	94.6	91.8	92.6	94.2	92.4	94.9
N.-É.	92.0	91.0	90.9	91.3	91.5	89.9
N.-B.	91.1	90.3	91.0	90.0	89.0	90.4
QC	91.6	92.2	92.6	93.2	91.4	91.6
Ont.	88.2	87.9	89.3	87.4	85.4	88.5
Man.	89.0	88.1	91.0	86.9	84.5	85.6
Sask.	86.1	86.2	89.3	86.4	87.2	87.2
Alb.	89.3	85.3	85.6	82.9	84.0	86.2
C.-B.	88.1	89.7	89.0	87.6	90.3	87.3
Canada	89.6	89.2	90.2	88.9	88.1	89.0

TABLEAU 7. **Taux de réponse mensuels selon la province, en
(suite) pourcentage, 1997**

	Juil.	Août	Sept	Oct.	Nov	Déc
T.-N.	92.7	92.8	89.9	89.0	88.7	90.1
Î.-P.-É.	94.5	92.8	90.5	89.6	89.5	89.4
N.-É.	91.6	92.2	90.4	87.9	88.9	87.2
N.-B.	91.7	90.8	89.8	85.8	89.6	85.6
QC	94.3	93.9	92.0	90.5	90.9	89.9
Ont.	89.5	90.0	85.7	85.8	86.3	79.8
Man.	89.4	90.7	91.8	90.1	89.7	85.2
Sask.	86.8	89.0	86.2	83.5	84.4	80.9
Alb.	86.1	86.1	85.0	81.5	81.5	79.8
C.-B.	90.5	89.4	87.1	85.5	88.2	82.8
Canada	90.6	90.8	88.4	86.9	87.7	84.0

Nota: Ces taux de réponse mensuels correspondant à la taille de l'échantillon de trois groupes de renouvellement.

***Effet du
plan de
sondage***

Le tableau suivant présente l'effet du plan de sondage, la taille de l'échantillon et les chiffres de population selon la province, qui ont servi à produire les tableaux des degrés approximatifs de variabilité de l'échantillonnage visant le poids de la personne. Il convient de noter qu'en dépit du fait que l'EVC se caractérise par des échantillons et des populations de taille différente chaque mois, l'effet du plan de sondage demeure constant d'un mois à l'autre. C'est pourquoi l'effet du plan de sondage, la taille de l'échantillon et l'effectif de la population ne sont présentés que pour un seul mois.

TABLEAU 8. *Effets du plan de sondage pour les trois groupes de renouvellement, décembre 1997*

Province	Effet du plan de sondage	Taille de l'échantillon	Population
T.-N.	1.99	843	448109
Î.-P.-É.	1.64	630	107049
N.-É.	1.87	1482	744970
N.-B.	1.74	1303	604583
QC	2.41	4444	5947756
Ont.	3.03	6171	9055307
Man.	1.92	1617	860889
Sask.	1.56	1521	762132
Alb.	1.78	1593	2193323
C.-B.	1.87	1956	3133986
Provinces atlantiques	1.99	4258	1904711
Man. et Sask.	1.92	3138	1623021
Alb. et C.-B.	1.87	3549	5327309
Canada	2.84	21560	23858104

**Seuils
relatifs à la
diffusion
pour l'EVC**

L'effectif minimum des groupes visés par l'estimation (si l'on tient compte des poids de la personne) est indiqué dans le tableau suivant pour le niveau provincial, régional et pour le Canada. Les estimations qui portent sur des effectifs inférieurs à ceux de l'effectif minimum indiqué dans la colonne Inacceptable ne peuvent être diffusées en aucun cas. Il convient de noter qu'un seul tableau affichant les seuils relatifs à la diffusion est présenté ci-dessous. Ce tableau donne les seuils relatifs à la diffusion pour décembre 1997.

TABLEAU 9. Spécimen de tableau des seuils relatifs à la diffusion, décembre 1997

Province	Acceptable (0.0-16.5)	Médiocre (16.6-25.0)	Confidentiel (25.1-33.3)	Inacceptable (33.4 & >)
T.-N.	36,000 +	16,500-35,999	9,500-16,499	moins de 9,500
Î.-P.-É.	9,500 +	4,500-9,499	2,500-4,499	moins de 2,500
N.-É.	33,000 +	14,500-32,999	8,500-14,499	moins de 8,500
N.-B.	28,500 +	12,500-28,499	7,000-12,499	moins de 7,000
QC	116,000 +	51,000-115,999	29,000-50,999	moins de 29,000
Ont.	160,500 +	70,500-160,499	40,000-70,499	moins de 40,000
Man.	36,000 +	16,000-35,999	9,000-15,999	moins de 9,000
Sask.	27,500 +	12,500-27,499	7,000-12,499	moins de 7,000
Alb.	86,500 +	38,500-86,499	22,000-38,499	moins de 22,000
C.-B.	106,500 +	47,000-106,499	27,000-46,999	moins de 27,000
Canada	115,000 +	50,000-114,999	28,500-49,999	moins de 28,500

**Obtention
des CV
approximatif
à partir des
tableaux**

Les coefficients de variation approximatifs (CV) sont présentés dans les tableaux des degrés approximatifs de variabilité de l'échantillonnage à la fin de ce chapitre et sur le CD-ROM. Avant d'appliquer le critère du coefficient de variation, il importe de suivre les lignes directrices fondées sur la taille des échantillons décrites au [chapitre 9](#) (Lignes directrices pour la diffusion).

Les règles et les exemples qui suivent devraient permettre à l'utilisateur de déterminer, à partir des tableaux de variabilité de l'échantillonnage, les coefficients de variation approximatifs d'estimations du nombre de personnes qui, dans la population observée, possèdent une caractéristique donnée. Les exemples concrets visent à aider l'utilisateur à appliquer ces règles. Ils sont fondés sur des variables nécessitant l'application du poids de la personne en vue de la production des estimations. Les règles et les exemples suivants sont des exemples de 1996. Les mêmes principes s'appliquent en 1997

Règle 1 : Estimations du nombre de personnes possédant une caractéristique donnée (agrégats)

Le coefficient de variation dépend uniquement de la taille de l'estimation elle-même. Sur le tableau de variabilité de l'échantillonnage correspondant à la région géographique appropriée, repérez le nombre estimé dans la colonne à l'extrême gauche de la table (intitulée «Estimation») et suivez les astérisques (le cas échéant) jusqu'au premier nombre. Ce nombre constitue le coefficient de variation approximatif.

Exemple illustrant la règle 1

Supposons qu'un utilisateur estime à 6 032 234 le nombre de personnes ayant fait au moins un voyage en mars 1996. Comment détermine-t-il le coefficient de variation de cette estimation?

- ▶ Reportez-vous à la table de CV pour le CANADA (Tableaux de variance échantillonnale approximative - Poids des personnes) à la fin du présent chapitre.
- ▶ Comme l'estimation de l'agrégat (6 032 234) n'apparaît pas dans la colonne de gauche (intitulée «Numérateur du pourcentage»), il faut utiliser le chiffre qui s'en rapproche le plus, soit 6 000 000.
- ▶ Le coefficient de variation de l'estimation de l'agrégat correspond au premier chiffre suivant les astérisques dans cette rangée, soit 2,3 %.
- ▶ Par conséquent, le coefficient de variation approximatif de cette estimation est égal à 2,3 %. Le résultat obtenu, soit que 6 032 234 personnes ont fait au moins un voyage en mars 1996, peut donc être publié sans restriction.

Règle 2 : Estimations de proportions ou de pourcentages de personnes possédant une caractéristique donnée

Le coefficient de variation de l'estimation d'une proportion ou d'un pourcentage dépend à la fois de la taille de la proportion ou du pourcentage et de la taille de la population totale à laquelle cette proportion ou ce pourcentage se rapporte. Les

estimations de proportions ou de pourcentages sont relativement plus fiables que les estimations correspondantes du numérateur de la proportion ou du pourcentage lorsque la proportion ou le pourcentage se rapporte à un sous-groupe de la population. Par exemple, la proportion des «personnes âgées de 15 ans et plus ayant fait au moins un voyage au cours du mois de référence» est plus fiable que l'estimation du nombre de «personnes âgées de 15 ans et plus ayant fait au moins un voyage au cours du mois de référence». (Il convient de noter que, dans les tableaux, la valeur des CV décroît de la gauche vers la droite).

Lorsque la proportion ou le pourcentage se rapporte à la population totale de la région géographique visée par le tableau, le CV de cette proportion ou de ce pourcentage est égal à celui du numérateur de la proportion ou du pourcentage. Dans ce cas, on peut utiliser la règle 1.

Lorsque la proportion ou le pourcentage se rapporte à un sous-ensemble de la population totale (p. ex., aux personnes d'un sexe ou d'un groupe d'âge donné), on doit se reporter à la proportion ou au pourcentage (rangée au haut du tableau) et au numérateur de cette proportion ou de ce pourcentage (colonne de gauche du tableau). Le coefficient de variation se trouve à l'intersection de la rangée et de la colonne appropriées.

Exemple illustrant la règle 2

Supposons qu'un utilisateur estime à $2\,951\,511 / 6\,032\,234 = 49\%$ le pourcentage des personnes qui, ayant voyagé en mars, ont effectué au moins un

voyage du même jour. Comment l'utilisateur détermine-t-il le coefficient de variation de cette estimation?

- ▶ Reportez-vous à la table visant le CANADA. (Tableaux de variance échantillonnale approximative - Poids des personnes) à la fin du présent chapitre.
- ▶ Étant donné que l'estimation est un pourcentage se rapportant à un sous-ensemble de la population totale (soit les voyageurs qui ont fait au moins un voyage du même jour en mars), il faut utiliser à la fois le pourcentage (49 %) et le numérateur de ce pourcentage (2 951 511) pour déterminer le coefficient de variation.
- ▶ Le numérateur, 2 951 511, n'apparaissant pas dans la colonne de gauche (intitulée «Numérateur du pourcentage»), il convient d'utiliser le chiffre qui s'en rapproche le plus, soit 3 000 000. De la même façon, l'estimation du pourcentage ne figure dans aucun des titres de colonne, et il faut donc se reporter au pourcentage qui s'en rapproche le plus, soit 50,0 %.
- ▶ Le chiffre indiqué à l'intersection de la rangée et de la colonne consultées, soit 2,8 %, correspond au coefficient de variation recherché.
- ▶ Par conséquent, le coefficient de variation approximatif de cette estimation est 2,8 %. Le résultat obtenu, soit que 49 % des personnes ayant voyagé en mars ont fait au moins un voyage du même jour, peut donc être publié sans restriction.

Règle 3 : Estimations des différences entre des agrégats ou des pourcentages

L'erreur-type de la différence entre deux estimations correspond à peu près à la racine carrée de la somme des carrés de chaque erreur-type considérée séparément. Par conséquent, l'erreur-type de la différence ($\bar{d} = X_1 - X_2$) se calcule ainsi :

$$\sigma_{\bar{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

où X_1 représente l'estimation 1, X_2 , l'estimation 2, et α_1 et α_2 sont les coefficients de variation de X_1 et de X_2 , respectivement. Le coefficient de variation de \bar{d} est donné par $\sigma_{\bar{d}}/\bar{d}$. Cette formule donne des résultats précis lorsqu'on l'applique à la différence entre des caractéristiques distinctes et non corrélées; elle donne des résultats approximatifs dans les autres cas.

Exemple illustrant la règle 3

Supposons qu'un utilisateur estime à $2\,951\,511 / 6\,032\,234 = 49\%$ le pourcentage de personnes qui, ayant voyagé en mars, ont fait au moins un voyage du même jour et à $3\,998\,785 / 6\,032\,234 = 66,3\%$ le pourcentage de personnes qui, ayant voyagé en mars, ont fait au moins un voyage d'une nuit ou plus. (Notons qu'une même personne peut avoir fait un voyage du même jour et un voyage d'une nuit ou plus au cours du même mois, ce qui se traduirait par un chevauchement des deux estimations). Comment l'utilisateur détermine-t-il le coefficient de variation de la différence entre ces deux estimations?

- ▶ L'utilisation de la table des CV pour le CANADA selon la méthode décrite à l'exemple 2 permet d'établir à 2,8 % le CV de l'estimation des voyageurs ayant fait au moins un voyage du même jour et à 2,1 % le CV des voyageurs ayant fait au moins un voyage d'une nuit ou plus.
- ▶ Conformément à la règle 3, on calcule l'erreur-type d'une différence exprimée sous la forme ($\bar{d} = X_1 - X_2$) comme suit :

$$\sigma_{\bar{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

où X_1 représente l'estimation 1, X_2 , l'estimation 2, et α_1 et α_2 sont les coefficients de variation de X_1 et de X_2 , respectivement.

L'erreur-type de la différence $\bar{d} = (.663 - .490) = .173$ est donc égale à :

$$\begin{aligned}\sigma_{\bar{d}} &= \sqrt{[(.49)(.028)]^2 + [(.660)(.021)]^2} \\ &= \sqrt{(.00001882384) + (.0001920996)} \\ &= .0195022\end{aligned}$$

Le coefficient de variation de \bar{d} correspond à $\sigma_{\bar{d}}/\bar{d} = .019/.173 = 0.11$.

- ▶ Par conséquent, le coefficient de variation approximatif de la différence entre les estimations est égal à 11 %. Cette estimation peut donc être publiée sans restriction.

Règle 4 : Estimations de rapports

Si le numérateur est un sous-ensemble du dénominateur, il faut exprimer le rapport en pourcentage et appliquer la règle 2. Cette situation se produirait, par exemple, si le dénominateur correspondait au nombre de «personnes ayant fait au moins un voyage au cours du mois de référence» et si le numérateur correspondait au nombre de «personnes ayant fait au moins un voyage d'affaires au cours du mois de référence».

Si le numérateur n'est pas un sous-ensemble du dénominateur (par exemple, s'il s'agit du rapport entre le nombre de «personnes ayant fait au moins un voyage d'affaires au cours du mois de référence» et le nombre de «personnes ayant fait au moins un voyage d'agrément au cours du mois de référence»), l'écart-type du rapport entre les estimations correspond à peu près à la racine carrée de la somme des carrés de chaque coefficient de variation considéré séparément multipliée par R. L'erreur-type d'un rapport exprimée sous la forme ($R = X_1 / X_2$) se calcule donc ainsi :

$$\sigma_{\hat{R}} = \hat{R} \sqrt{\alpha_1^2 + \alpha_2^2}$$

où α_1 et α_2 représentent les coefficients de variation de X_1 et de X_2 respectivement.

Le coefficient de variation de R est donné par σ_R/R . La formule tendra à surestimer l'erreur s'il existe une corrélation positive entre X_1 et X_2 et à sous-estimer l'erreur s'il existe une corrélation négative entre X_1 et X_2 .

Exemple illustrant la règle 4

Supposons qu'un utilisateur estime à 3 998 785 le nombre de personnes qui, ayant voyagé en mars, ont fait au moins un voyage d'une nuit ou plus et à 2 951 511 le nombre de personnes qui, ayant voyagé en mars, ont fait au moins un voyage du même jour. L'utilisateur cherche à comparer l'estimation du nombre de voyageurs ayant fait des voyages d'une nuit ou plus à celle du nombre de voyageurs ayant fait des voyages du même jour et veut exprimer cette comparaison sous la forme d'un rapport. Comment l'utilisateur détermine-t-il le coefficient de variation de cette estimation?

- ▶ Tout d'abord, cette estimation est une estimation par ratio dont le numérateur ($= \bar{X}_1$) correspond au nombre de personnes qui, ayant voyagé en mars, ont fait au moins un voyage d'une nuit ou plus. Le dénominateur de l'estimation ($= \bar{X}_2$) correspond au nombre de personnes qui, ayant voyagé en mars, ont fait au moins un voyage du même jour.
- ▶ Reportez-vous à la table visant le CANADA. (Tableaux de variance échantillonnale approximative - Poids des personnes) à la fin du présent chapitre.
- ▶ Le numérateur de cette estimation par ratio est 3 998 785. Le chiffre qui s'en rapproche le plus est 4 000 000. Le coefficient de variation de cette estimation correspond au premier chiffre suivant les astérisques dans cette rangée, 3,1%.

- ▶ Le dénominateur de cette estimation par ratio est 2 951 511. Le chiffre qui s'en rapproche le plus est 3 000 000. Le coefficient de variation de cette estimation correspond au premier chiffre suivant les astérisques dans cette rangée, 3,6%.
- ▶ Conformément à la règle 4, le coefficient de variation approximatif de cette estimation par ratio se calcule comme suit :

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^2 + \alpha_2^2}$$

où α_1 et α_2 représentent les coefficients de variation de X_1 et de X_2 respectivement.

Par conséquent :

$$\begin{aligned}\alpha_{\hat{R}} &= \sqrt{(.031)^2 + (.036)^2} \\ &= 0.047\end{aligned}$$

- ▶ Le rapport obtenu entre les voyageurs de mars 1996 ayant fait au moins un voyage d'une nuit ou plus et les voyageurs de mars 1996 ayant fait au moins un voyage du même jour est de 3 998 785 / 2 951 511, soit 1.35:1. Le coefficient de variation de cette estimation s'établit à 4,7%, et l'estimation peut donc être diffusée sans restriction.

Règle 5 : Estimations de différences entre des rapports

Dans ce cas, on combine les règles 3 et 4. On calcule d'abord les CV des deux rapports au moyen de la règle 4, puis le CV de leur différence

conformément à la règle 3.

**Utilisation
des tables
de CV pour
déterminer
les
intervalles
de
confiance**

Bien que les coefficients de variation soient largement utilisés, on peut mesurer de façon significative l'erreur d'échantillonnage par une mesure plus intuitive, soit l'intervalle de confiance d'une estimation. L'intervalle de confiance constitue une assertion quant à la probabilité de confiance que la valeur réelle de la population se situe dans une fourchette de valeurs donnée. Par exemple, un intervalle de confiance de 95 % peut se définir comme suit :

Si l'échantillonnage de la population se répète à l'infini, chacun des échantillons modifiant l'intervalle de confiance d'une estimation, l'intervalle contiendra la véritable valeur de la population dans 95 % des cas.

Une fois que l'erreur-type d'une estimation est déterminée, on peut calculer les intervalles de confiance d'une estimation en partant de l'hypothèse que les diverses estimations d'une caractéristique donnée d'une population, obtenues à partir d'un échantillonnage répété de cette population, se dispersent normalement autour de la valeur réelle de la population. Compte tenu de cette hypothèse, dans environ 68 cas sur 100, l'écart entre une estimation par échantillonnage et la véritable valeur de la population sera inférieur à une erreur-type, dans environ 95 cas sur 100, cet écart sera inférieur à deux erreurs-types et, dans environ 99 cas sur 100, l'écart sera inférieur à trois erreurs-types. On appelle ces différents degrés de confiance des niveaux de confiance.

Les intervalles de confiance pour une estimation \hat{X} sont généralement exprimés sous la forme de deux nombres, l'un étant inférieur à l'estimation et l'autre, supérieur à celle-ci, sous la forme ($\hat{X}-k$, $\hat{X}+k$) où k varie selon le niveau de confiance désiré et l'erreur d'échantillonnage de l'estimation.

On peut calculer directement les intervalles de confiance d'une estimation à partir des tableaux des degrés approximatifs de variabilité de l'échantillonnage en déterminant d'abord, au moyen du tableau approprié, le coefficient de variation de l'estimation \hat{X} , puis en utilisant la formule suivante pour obtenir l'intervalle de confiance CI correspondant :

$$CI_X = [\hat{X} - t\hat{X}\alpha_{\hat{X}}, \hat{X} + t\hat{X}\alpha_{\hat{X}}]$$

où $\alpha_{\hat{X}}$ est le coefficient de variation obtenu pour \hat{X} ,

$t = 1$ si l'on désire un intervalle de confiance de 68 %
 $t = 1.6$ si l'on désire un intervalle de confiance de 90 %
 $t = 2$ si l'on désire un intervalle de confiance de 95 %
 $t = 3$ si l'on désire un intervalle de confiance de 99 %.

Note : Les lignes directrices qui s'appliquent à la diffusion des estimations s'appliquent aussi aux intervalles de confiance. Par exemple, si l'estimation ne peut être diffusée, alors l'intervalle de confiance ne peut, lui non plus, être diffusé.

**Exemple
d'utilisation
des tables
de CV pour
déterminer
les
intervalles
de
confiance**

Un intervalle de confiance de 95 % pour l'estimation de la proportion des personnes qui ont voyagé en mars et qui ont fait au moins un voyage du même jour (tiré de l'exemple illustrant la [règle 2](#) du chapitre 10) se calculerait comme suit.

$$\hat{X} = 49 \% \text{ (ou, sous forme de proportion, } = .49)$$

$$t = 2$$

$\alpha_X = 2,8 \%$ (soit la proportion .028 exprimée en pourcentage) est le coefficient de variation de cette estimation déterminé à partir des tables.

$$CI_X = \{.49 - (2) (.49) (.028), .49 + (2) (.49) (.028)\}$$

$$CI_X = \{.49 - .027, .49 + .027\}$$

$$CI_X = \{.463, .517\}$$

On peut affirmer avec un degré de confiance de 95 % qu'entre 46,3 % et 51,7 % des personnes qui ont voyagé en mars ont fait au moins un voyage du même jour.

**Utilisation
des tables
de CV pour
effectuer un
test t**

On peut aussi utiliser les erreurs-types pour effectuer des tests d'hypothèses, une technique qui permet de faire la distinction entre les paramètres d'une population à l'aide d'estimations fondées sur un échantillon. Ces estimations peuvent être des nombres, des moyennes, des pourcentages, des rapports, etc. Les tests peuvent être effectués à divers niveaux de signification, le niveau de signification étant la probabilité de conclure que les

caractéristiques sont différentes quand, en fait, elles sont identiques.

Soit \mathbf{X}_1 et \mathbf{X}_2 des estimations fondées sur un échantillon visant deux caractéristiques données. Supposons que l'erreur-type de la différence $\mathbf{X}_1 - \mathbf{X}_2$ est σ_d .

Si $t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_d}$ se situe entre -2 et 2, alors on ne

peut tirer aucune conclusion quant à la différence entre les caractéristiques à un seuil de signification de 5 %. En revanche, si ce rapport est inférieur à -2 ou supérieur à +2, la différence observée est significative considérant un seuil de 0.05. En d'autres mots, les caractéristiques sont significatives.

Exemple d'utilisation des tables de CV pour effectuer un test t

Supposons que l'on désire tester, à un niveau de signification de 5 %, l'hypothèse qu'il n'existe pas de différence entre la proportion des voyageurs en mars qui ont fait au moins un voyage du même jour et la proportion de voyageurs en mars qui ont fait au moins un voyage d'une nuit ou plus. Dans l'exemple illustrant la règle 3, l'erreur-type de la différence entre ces deux estimations a été établie à .019. Par conséquent,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_d} = \frac{.49 - .663}{.019} = \frac{-.173}{.019} = -9.10.$$

Puisque $t = -9.10$ est inférieur à -2, on doit conclure qu'il existe une différence significative entre les deux estimations à un seuil de signification de 0.05.

CV des estimations quantitatives

Pour calculer l'erreur d'échantillonnage d'estimations quantitatives, il faudrait élaborer des tableaux spéciaux, ce qui n'a pas été fait puisque la plupart des variables de l'Enquête sur les voyages des Canadiens sont essentiellement de type nominal. Les tableaux sont compris dans la documentation et peuvent être utilisés dans le cadre de l'analyse des variables qui nécessitent l'application des poids du voyage-personne, du voyage-ménage, de la nuitée-personne et des dépenses.

En règle générale, le coefficient de variation d'un total quantitatif est supérieur à celui de l'estimation nominale correspondante (c.-à-d, l'estimation du nombre de personnes contribuant à l'estimation quantitative). Si l'estimation nominale correspondante ne peut être diffusée, il en sera de même pour l'estimation quantitative. Par exemple, le coefficient de variation du nombre total de voyages effectués en mars sera supérieur au coefficient de variation de la proportion correspondante des personnes qui ont fait au moins un voyage en mars. Par conséquent, si le coefficient de variation de la proportion ne peut être diffusé, il en sera de même pour le coefficient de variation de l'estimation quantitative correspondante.

Dans le cas d'estimations de ce genre, on peut déterminer au besoin les coefficients de variation d'une estimation particulière à l'aide d'une technique appelée pseudo-réplication qui consiste à diviser les enregistrements des fichiers de microdonnées en sous-groupes (ou sous-

échantillons) et à calculer la variabilité de l'estimation d'un sous-échantillon à l'autre. Les utilisateurs qui désirent calculer les coefficients de variation d'estimations quantitatives peuvent demander conseil à Statistique Canada relativement à la répartition des enregistrements en sous-échantillons appropriés et aux formules à utiliser pour ces calculs.

**Utilisation
pratique des
tableaux de
CV se
rapportant
aux voyages**

Structure des tableaux de CV

La section de ce document intitulée Tableaux de CV fournit aux utilisateurs une mesure qualitative des estimations de l'Enquête sur les voyages des canadiens. Ces tableaux sont produits pour le Canada, les provinces et les régions et ce pour les estimations mensuelles, trimestrielles ou annuelles. Il s'agit donc d'un ensemble de 84 tableaux répartis en deux ensembles distincts soit: 42 tableaux de CV estimatifs pour les voyages de type intraprovinciaux et totaux et 42 autres pour les voyages interprovinciaux. On doit noter qu'en 1997 le nombre de tableaux de CV a doublé par rapport à 1996 étant donné que pour cette année deux différentes tailles d'échantillons ont été utilisées dépendant qu'il s'agisse de voyages faits à l'intérieur ou à l'extérieur d'une même province. Pour plus de détails se rapportant à l'échantillon voir le [Chapitre 3](#) de ce guide.

Utilisation des tableaux de CV au niveau provincial et infraprovincial

Si on veut connaître les coefficients de variation d'estimations se rapportant à une province précise ou à une entité géographique (région touristique, division de recensement, région métropolitaine de recensement ou agglomération de recensement) de cette province, on utilise les CV provenant des tableaux identifiés à cette province.

Lorsque l'utilisateur est en présence de tableaux croisés dans lesquels l'origine et la destination des voyages se retrouvent dans des provinces différentes, les CV de la province de destination doivent être utilisés.

Panoplie de coefficients de variation

Les coefficients de variation apparaissant dans la série de tableaux de CV se rapportent aux unités de mesure suivantes: les voyages-ménages, les voyages-personnes, les nuitées-personnes et les dépenses-ménages. Lorsque les estimations produites ont comme unité de mesure les visites-personnes, on utilise les CV se rapportant aux voyages-personnes et lorsqu'elles sont en nuitées-visites-personnes, on utilise ceux des nuitées-personnes.

Comment trouver le bon coefficient à partir des tableaux de CV de la section de ce GUIDE intitulée TABLEAUX de CV

Exemple numéro 1

On veut obtenir le coefficient de variation approximatif d'une estimation *annuelle* de 10 000 000 de voyages-personnes ou visites-personnes, comprenant l'ensemble des voyages effectués en *Ontario* en 1997.

- ▶ On se rend à la section du guide intitulée:
Tableaux de CV
- ▶ On trouve le tableau de l'Ontario, ANNUEL,
VOYAGES TOTAL & INTRAPROVINCIAUX.
- ▶ Au niveau de 10 000 000 de l'échelle de gauche ,
on y associe le CV correspondant de la colonne
voyages-personnes, lequel correspond à 4,8%.

Si le CV recherché dans cet exemple avait été pour une valeur de 10 000 000 de dollars au lieu de 10 000 000 des voyages-personnes, on aurait pris la lecture du CV en se servant cette fois de l'échelle d'estimation de droite. La lecture du CV correspondant aurait été de 34,0%.

Supposons que la valeur de l'exemple numéro 1 est de 12 000 000 de voyages-personnes au lieu de 10 000 000 de voyages-personnes. Comme cette nouvelle valeur n'apparaît pas directement au tableau, voici comment en dériver le CV.

Méthode rapide: utiliser la valeur du CV la plus près de la valeur recherchée. Ici le CV le plus près de 12 000 000 est celui associé à 10 000 000 soit 4,8%.

Méthode précise: utiliser la formule qui suit.

$$A = CV1 \quad B = CV2$$

$$CV_X = (EST_X - A) / (B - A) * (CV1 - CV2) + CV2$$

10 000 000 inférieure CV = 4,8

15 000 000 supérieure CV = 4,2

et vous voulez associer un CV à une valeur de 12 000 000 :

$$CV_X = (12\,000\,000 - 10\,000\,000) / (15\,000\,000 - 10\,000\,000) * (4,8 - 4,2) + 4,2 = 4,44\%$$

Exemple numéro 2:

On veut obtenir le coefficient de variation approximatif d'une estimation *trimestrielle* de 10 000 000 de voyages-ménages comprenant seulement les voyages intraprovinciaux effectués en *Ontario* en 1997 pour un trimestre donné.

- ▶ On se rend à la section du guide intitulée:
Tableaux de CV
- ▶ On trouve le tableau de l'Ontario,
TRIMESTRIELLE, VOYAGES TOTAL &
INTRAPROVINCIAUX
- ▶ Au niveau de 10 000 000 de l'échelle de gauche ,
on y associe le CV correspondant de la colonne
voyages-ménages, lequel correspond à 5,9%.

Si l'estimation pour laquelle on recherche un CV n'apparaît pas au tableau, on peut utiliser l'une des deux méthodes indiquées à l'exemple numéro 1.

Exemple numéro 3:

On veut obtenir le coefficient de variation approximatif d'une estimation *mensuelle* de 10 000 000 de nuitées-personnes pour les voyages interprovinciaux effectués en *Ontario* en 1997 pour un mois donné.

- ▶ On se rend à la section du guide intitulée:
Tableaux de CV
- ▶ On trouve le tableau de l'Ontario, MENSUEL, VOYAGES INTERPROVINCIAUX.
- ▶ Au niveau de 10 000 000 de l'échelle de gauche, on y associe le CV correspondant de la colonne nuitées-personnes, lequel correspond à 8,2%.

Si l'estimation pour laquelle on recherche un CV n'apparaît pas au tableau, on peut utiliser l'une des deux méthodes indiquées à l'exemple numéro 1.

