

ANNEXE 34 – Précision des données

Ce paragraphe reprend les principaux résultats d'une étude réalisée par le Cetur en 1977 et actualisée par le Certu à laquelle le lecteur pourra se référer pour plus de détails²¹.

L'échantillonnage des enquêtes ménages déplacements étant réalisé selon une méthode statistique aléatoire, il est possible d'appliquer les lois de la statistique pour évaluer la précision des résultats.

En outre, l'une des particularités des enquêtes ménages déplacements réside dans le mode de tirage « en grappe » des personnes et des déplacements : toutes les personnes des ménages tirés au sort sont enquêtées sur leurs déplacements. Les échantillons de personnes et de déplacements qui en résultent ne sont plus totalement aléatoires : il est nécessaire d'intégrer ce que l'on appelle « l'effet de grappe » dans l'estimation de la précision des résultats.

Comme le taux de sondage est faible, la fiabilité statistique des résultats issus de l'enquête ménages déplacements dépend essentiellement des effectifs enquêtés au sein de l'échantillon (avant redressement). Ainsi, pour un critère d'analyse donné, plus le nombre d'individus enquêtés est grand, plus la précision du résultat apporté par l'enquête est grande.

1 - CALCULS DE PRÉCISION

Les calculs de précision sont différents selon qu'il s'agit de proportions (comme les parts de marché ou les matrices « origine-destination ») ou de moyennes.

1.1 - Proportions et matrices « origine-destination »

Le tableau suivant indique les incertitudes d'estimation relatives au calcul de proportions et de matrices « origine-destination » lors de l'exploitation d'enquêtes ménages déplacements. Les formules indiquées tiennent compte de l'effet de grappe²² et varient donc, selon que le calcul porte sur les ménages, les personnes ou les déplacements. Les formules contenues dans ce tableau permettent de calculer l'intervalle dans lequel se trouve la valeur réelle d'une proportion, au seuil de confiance de 95 % (c'est-à-dire que la valeur réelle a 95 % de chances de se trouver dans l'intervalle calculé).

Les intervalles donnés s'appliquent aux proportions calculées à partir des valeurs brutes enquêtées.

Dans le tableau ci-après, « n » est l'effectif total brut enquêté pour le calcul de cette proportion ; « p » est la proportion calculée à partir des effectifs bruts de l'enquête (exemple : répartition de la population par PCS, répartition des déplacements par mode, par motif, etc.); la proportion réelle (p réel) est comprise dans l'intervalle qui est défini dans le tableau.

²¹ *Éléments pour évaluer la précision obtenue dans l'estimation des indicateurs donnés par les enquêtes ménages*, Certu, décembre 2006.

²² On a pris un effet de grappe de 1,2 pour les personnes et de 2 pour les déplacements, chiffres couramment observés. En pratique, l'effet de grappe est différent selon les variables analysées et doit être calculé au cas par cas.

Échantillon concerné	Proportions
	Intervalle de confiance au seuil de confiance à 95 % ²³
Ménages	$p \text{ réel} \in p \pm 1,96 \sqrt{\frac{p(1-p)}{n}}$
Personnes	$p \text{ réel} \in p \pm 2,35 \sqrt{\frac{p(1-p)}{n}}$
Déplacements	$p \text{ réel} \in p \pm 3,92 \sqrt{\frac{p(1-p)}{n}}$

Remarque : la formule donnée par le tableau n'est applicable que si n est suffisamment grand ($n > 30$) et que si np et $n(1-p) > 5$; pour plus de précision, et dans les cas où les conditions ci-dessus ne seraient pas remplies, on se reportera, pour le calcul de l'intervalle de confiance, aux abaques figurant dans l'étude du Cetur citée plus haut.

Le cas d'une matrice « origine-destination » est un cas particulier de proportion : si n représente l'effectif total brut de la matrice, n_{ij} l'effectif brut de la liaison $i \rightarrow j$, $p_{ij} = n_{ij}/n$ la proportion calculée de la liaison $i \rightarrow j$ dans le total des déplacements, l'intervalle dans lequel se trouve la valeur réelle de p_{ij} (p_{ij} réel) se calcule à partir de la formule située à la ligne « déplacements »²⁴.

En pratique, dans le cas de matrices « origine-destination », on observe fréquemment que, si le nombre brut de déplacements enquêtés sur la liaison $i \rightarrow j$ est inférieur à 10, la précision devient si mauvaise qu'il faut éviter d'utiliser ce résultat (l'erreur relative dépasse en effet alors 100 %). Il faut alors effectuer des regroupements de plusieurs liaisons.

Précision obtenue sur les parts de marché des modes de déplacements en fonction du nombre de ménages enquêtés (effet de grappe²⁵ compris)

	Valeur observée	Nombre de ménages enquêtés					
		15 000	10 000	5 000	3 000	2 000	1 000
Part de marché de la voiture	56 %	± 0,5 %	± 0,7 %	± 0,9 %	± 1,2 %	± 1,5 %	± 2,1 %
Part de marché des transports collectifs	10 %	± 0,3 %	± 0,4 %	± 0,6 %	± 0,7 %	± 0,9 %	± 1,3 %
Part de marché des deux-roues	5 %	± 0,2 %	± 0,3 %	± 0,4 %	± 0,5 %	± 0,6 %	± 0,9 %
Part de marché de la marche à pied	29 %	± 0,5 %	± 0,6 %	± 0,9 %	± 1,1 %	± 1,3 %	± 1,9 %

Guide de lecture : au seuil de confiance de 95 % (c'est-à-dire avec 1 risque sur 20 de se tromper), la part de marché de la voiture est comprise entre 54,5 % et 57,5 % si l'on a enquêté 2 000 ménages.

²³ On a pris un effet de grappe de 1,2 pour les personnes et de 2 pour les déplacements, chiffres couramment observés. En pratique, l'effet de grappe est différent selon les variables analysées et doit être calculé au cas par cas.

²⁴ En toute rigueur, il faudrait ajouter à l'imprécision sur la proportion, l'imprécision sur le total des déplacements. Mais cette dernière imprécision est généralement négligeable par rapport à celle sur la proportion.

²⁵ Id 24.

1.2 - Moyennes

Il n'existe pas de formule simple permettant de calculer l'incertitude sur l'évaluation de moyennes. Les tableaux ci-après indiquent toutefois la précision obtenue sur les mobilités moyennes ainsi que les effectifs minima bruts pour obtenir différentes précisions dans le calcul de moyennes de certaines variables.

Précision obtenue sur les mobilités moyennes selon les modes de déplacements en fonction du nombre de ménages enquêtés (effet de grappe²⁶ compris)

	Valeur observée	Nombre de ménages enquêtés					
		15 000	10 000	5 000	3 000	2 000	1 000
Taille moyenne des ménages	2,50	± 0,03	± 0,03	± 0,04	± 0,06	± 0,07	± 0,10
Nombre de voitures par ménage	1,00	± 0,01	± 0,02	± 0,03	± 0,03	± 0,04	± 0,06
Mobilité tous modes	3,50	± 0,04	± 0,04	± 0,06	± 0,08	± 0,10	± 0,14
Mobilité en voiture	2,00	± 0,03	± 0,04	± 0,06	± 0,08	± 0,09	± 0,13
Mobilité en transports collectifs	0,30	± 0,01	± 0,01	± 0,02	± 0,02	± 0,03	± 0,04
Mobilité en deux-roues	0,10	± 0,01	± 0,01	± 0,01	± 0,02	± 0,02	± 0,03
Mobilité à pied	1,10	± 0,02	± 0,03	± 0,04	± 0,05	± 0,06	± 0,09

Guide de lecture : au seuil de confiance de 95 % (c'est-à-dire avec 1 risque sur 20 de se tromper), la mobilité moyenne en voiture est comprise entre 1,91 et 2,09 si l'on a enquêté 2 000 ménages.

²⁶ On a pris un effet de grappe de 1,2 pour les personnes et de 2 pour les déplacements, chiffres couramment observés. En pratique, l'effet de grappe est différent selon les variables analysées et devrait être calculé au cas par cas.

Variable	Effectifs minima bruts à atteindre Précision au seuil de confiance 95 %						Unité
	1 %	5 %	10 %	20 %	30 %	50 %	
Taille moyenne des ménages	13 800	550	140	35	15	6	ménages
Nombre d'actifs par ménage	27 800	1 100	280	70	30	10	ménages
Nombre de voitures par ménage	38 500	1 550	385	95	43	15	ménages
Taux d'occupation des voitures (tous motifs)	48 000	1 920	480	120	52	20	déplacements en tant que conducteur de voiture
Mobilité tous modes	30 000	1 200	300	70	33	12	personnes
Mobilité en modes mécanisés	79 000	3 200	800	200	85	30	personnes
Mobilité en voiture particulière	133 000	5 400	1 300	300	150	53	personnes
Mobilité en transports collectifs	390 000	15 600	3 900	980	440	160	personnes
Mobilité en deux-roues	707 000	28 000	7 100	1 750	800	280	personnes
Mobilité à pied	91 500	3 600	900	230	100	40	personnes

Guide de lecture : pour connaître à 5 % près au seuil de confiance de 95 % (c'est-à-dire avec 1 risque sur 20 de se tromper) la mobilité en voiture particulière, il faut disposer d'un effectif minimal de 5 400 personnes.

1.3 – Évolutions

Si l'on veut mesurer une évolution, plus elle sera petite, plus l'échantillon devra être grand pour qu'elle soit significative.

Premier exemple :

- une première enquête est réalisée auprès d'un échantillon de 2 300 ménages, soit 5 290 personnes (taille moyenne des ménages : 2,3 personnes) ayant réalisé au total 19 890 déplacements (mobilité moyenne : 3,76 déplacements par jour et par personne). La part de marché des transports collectifs obtenue est de 7 %. On a alors 95 chances sur 100 pour que la vraie valeur soit dans l'intervalle [6,65 % - 7,35 %] appelé « intervalle de confiance à 95 % » ;
- une seconde enquête est réalisée auprès d'un échantillon de 2 300 ménages soit 4 830 personnes (taille moyenne des ménages en légère baisse : 2,2 personnes) ayant réalisé 18 547 déplacements (mobilité moyenne : 3,84 déplacements par jour et par personne). La part de marché des transports collectifs obtenue est de 8 %. On a alors 95 chances sur 100 pour que la vraie valeur soit dans l'intervalle [7,61 % - 8,39 %].

Les deux intervalles de confiance étant disjoints, l'évolution de la part de marché des transports collectifs est bien significative.

Second exemple :

- une première enquête est réalisée auprès d'un échantillon de 2 300 ménages, soit 5 290 personnes (taille moyenne des ménages : 2,3 personnes) ayant réalisé au total 19 890 déplacements (mobilité moyenne : 3,76 déplacements par jour et par personne). La part de marché des transports collectifs obtenue est de 7 %. On a alors 95 chances sur 100 pour que la vraie valeur soit dans l'intervalle [6,65 % - 7,35 %] appelé « intervalle de confiance à 95 % » ;
- une seconde enquête est réalisée auprès d'un échantillon de 2 300 ménages soit 4 830 personnes (taille moyenne des ménages en légère baisse : 2,2 personnes) ayant réalisé 18 547 déplacements (mobilité moyenne : 3,84 déplacements par jour et par personne). La part de marché des transports collectifs obtenue est de 7,5 %. On a alors 95 chances sur 100 pour que la vraie valeur soit dans l'intervalle [7,12 % - 7,38 %].

Malgré les apparences, on ne pourra pas en conclure que les transports collectifs ont gagné un demi point de part de marché. En effet, il se peut que les vraies valeurs des parts de marché – mais par définition, on ne les connaît pas – soient, par exemple, de 7,3 % la première année et de 7,2 % la seconde année (ces valeurs sont bien comprises dans les intervalles de confiance respectifs des estimations obtenues). On pourra donc seulement conclure à la **quasi-stabilité** de la part de marché des transports collectifs.

Pour plus d'information, voir le document « *Éléments pour évaluer la précision obtenue dans l'estimation des indicateurs donnés par les enquêtes ménages* », Certu, décembre 2006.