Techniques d'enquête, décembre 1995
Vol. 21, n° 2, pp. 131–139
Statistique Canada

Effets du plan de sondage sur les $(P_i - P_j)$ corrélés

LESLIE KISH, MARTIN R. FRANKEL, VIJAY VERMA et NIKO KAĆIROTI[1]

RÉSUMÉ

En nous fondant sur 14 enquêtes menées dans six pays, nous présentons la constatation empirique de l'existence et de l'ampleur des effets du plan de sondage (eps) pour cinq plans appartenant à deux types principaux. Le premier type a trait à eps$(p_i - p_j)$, la différence de deux proportions d'une variable polytomique de trois catégories ou plus. Le deuxième type utilise les tests de chi carré pour l'analyse des différences entre deux échantillons. Nous montrons que pour toutes les variables et pour tous les plans, eps$(p_i - p_j) \cong$ [eps(p_i) + eps(p_j)]/2 constituent de bonnes approximations. Ces résultats sont *empiriques*, et les exceptions prouvent qu'il ne peut s'agir de simples inégalités analytiques. Il convient de signaler que ces résultats restent valables malgré les grandes variations des valeurs d'eps entre les variables et entre les catégories d'une même variable. Ils montrent en outre la nécessité d'utiliser des méthodes de traitement adaptées aux échantillons d'enquêtes pour l'analyse des données d'enquête, même lorsqu'on a affaire à des statistiques analytiques. En outre, ils permettent d'utiliser des approximations d'eps$(p_i - p_j)$ tirées des valeurs plus facilement accessibles d'eps(p_i).

MOTS CLÉS: Effets du plan de sondage; échantillonnage d'enquête; erreurs d'échantillonnage.

1. EFFETS DU PLAN DE SONDAGE SUR LES STATISTIQUES ANALYTIQUES

Nous nous penchons sur l'existence et sur l'ampleur des effets du plan de sondage sur certaines statistiques analytiques spéciales, en nous servant de données tirées d'échantillons d'enquêtes. Notre étude s'inspire à la fois de la méthode et de l'empirisme; elle s'appuie sur des données venant de plusieurs enquêtes aux variables différentes et provenant de populations contrastantes, et s'expose donc aux risques que présentent les résultats empiriques incohérents. On dit et on écrit souvent que l'échantillonnage probabiliste, bien que nécessaire à la réalisation d'enquêtes descriptives, n'est pas nécessaire pour les enquêtes analytiques. Dans une section intitulée "Four Obstacles to Representation in Analytic Studies", l'un de nous écrit qu'en plus des quatre obstacles concrets à la représentation auxquels il est fait allusion, il en existe un autre plus artificiel: le refus d'admettre la nécessité de la représentation (Kish 1987, section 2.7). Les études sur l'échantillonnage montrent que les méthodes de sélection probabiliste complexes, en particulier l'échantillonnage par grappes, n'ont pas d'effet appréciable sur les statistiques descriptives (p. ex., les moyennes et les coefficients de régression), mais qu'elles peuvent influer d'une façon extrêmement marquée sur les statistiques déductives comme les intervalles de confiance et les tests de signification (Kish et Frankel 1974).

Les effets du plan de sondage (eps) sont définis par la formule eps^2 = variance réelle/variance aléatoire simple, pour une même taille d'échantillon n (les deux variances étant estimatives). On a observé des valeurs d'eps supérieures à 1 pour des erreurs d'échantillonnage concernant non seulement les moyennes, mais également des statistiques analytiques comme les écarts de moyennes (et les tests de chi carré), les coefficients de régression, *etc.* Il est vrai qu'on a observé des réductions et des différences considérables des valeurs d'eps pour certaines statistiques analytiques. Les variations des valeurs d'eps ne sont pas une simple conséquence mathématique inévitable du choix du plan de sondage que l'on peut déduire une fois pour toutes. Elles ont un contenu empirique et doivent donc être reproductibles dans le cadre d'études empiriques (Kish et Frankel 1974; Kish 1987, 7.1; Kish 1965, 14.1-14.2; Rao et Wu 1985; Scott et Holt 1982; Skinner, Holt et Smith 1989). Dans le présent article, nous examinons l'incidence possible et l'importance des effets du plan de sondage sur un ensemble de statistiques apparentées qui n'ont fait l'objet d'aucune étude antérieure. Il se trouve que, dans de nombreux articles, des auteurs reconnus à bon droit ont simplement présumé que le choix du plan de sondage était sans effet, sans fournir aux lecteurs les mises en garde appropriées, et que cela n'a soulevé aucune objection de la part des examinateurs. Nous vérifierons donc si les eps sont réduits ou éliminés pour cet ensemble de statistiques analytiques (Cochran 1950; Mosteller 1952; Scott et Seber 1983; Seber et Wild 1993).

Par ailleurs, nous proposerons explicitement ce à quoi nous avons déjà fait allusion auparavant, à savoir que l'existence d'effets de sondage importants justifie largement le recours à la sélection probabiliste. Il serait difficile d'imaginer un modèle de distribution de population où le plan de sondage serait sans importance (ou sans intérêt) tout en produisant des effets considérables. Toutefois, à l'inverse, l'absence d'eps est nécessaire, quoique non suffisante, pour justifier la décision de ne pas recourir à la sélection probabiliste. Cette proposition donne du poids à notre étude qui met en rapport les valeurs d'eps$(p_i - p_j)$ des statistiques analytiques d'une part, et celles d'eps(p_i) et d'eps(p_j) pour deux des nombreuses catégories de la même variable d'autre part.

[1] Leslie Kish, ISR, University of Michigan, Ann Arbor MI 48106, U.S.A.; Martin R. Frankel, NORC et City University of New York; Vijay Verma, University of Essex, Colchester, C04 3SQ, U.K.; Niko Kaćiroti, Institut de statistiques, Tirana, Albanie.

Dans la section 2, nous décrivons les cinq problèmes apparentés (plans de sondage) au sujet desquels nous décrivons les erreurs d'échantillonnage à la section 3. Dans la section 4, nous examinons la constatation empirique exposée dans les tableaux. Dans la section 5, nous plaçons nos résultats dans le contexte des travaux antérieurs portant sur les valeurs d'eps pour les sous-classes et leurs différences.

2. STATISTIQUES SEMBLABLES POUR LES CINQ PLANS DE SONDAGE

On a montré que cinq plans de sondage (problèmes) appartenant à deux types distincts peuvent être traités à l'aide des mêmes statistiques simples (Kish 1965, Section 12.10). Pour notre présentation à la fois empirique et simple, nous utilisons des symboles qui dénotent les valeurs d'échantillons (eps, p_i et n_i), même si à l'occasion, l'usage de majuscules pour les valeurs de la population totale serait plus approprié.

La différence de proportions $p_2 - p_0 = n_2/n - n_0/n$ dénote l'estimation souhaitée, où $n = n_0 + n_1 + n_2 + \ldots n_k$ désigne la taille de l'échantillon, avec n unités choisies et pondérées également. Par ailleurs, en vertu des hypothèses de l'échantillonnage aléatoire simple, la variance de $(p_2 - p_0)$ est $(1 - f)[p_2 + p_0 - (p_2 - p_0)^2]/(n - 1)$.

Comparaisons de type A

1. La différence entre deux catégories $(n_2 - n_0)/n = (p_2 - p_0)$ d'une polytomie peut représenter la préférence manifestée pour deux partis parmi plusieurs (k) dans un sondage sur les intentions de vote, pour deux marques d'automobiles dans une étude de marchés, pour deux types d'attitudes, d'opinions ou de comportements concernant une variable quelconque, *etc.* Les autres choix $(k - 2)$ sont pris en compte dans p_1 et sont ignorés dans l'évaluation de la différence. (Voir à ce propos Scott et Seber 1983)

2. On peut attribuer les valeurs de rang de $-1, 0, +1$ (ou $0, 1, 2$ ou $c, c + 1, c + 2$) à une variable trichotomique ordonnée sans métrique, et les assimiler à une forme simple de la différence entre deux catégories. Cette forme est particulièrement utile pour le calcul des erreurs d'échantillonnage puisque les cinq plans de sondage peuvent par exemple utiliser $-1, 0, +1$ à titre de variable de calcul transformée.

3. La différence de proportions de deux variables différentes (x et y) peut être traitée comme en (1) et en (2). Ne définir comme positifs en x (succès) que les éléments qui sont positifs en x, mais non en y, de manière que $n_{10} = n(x_1, y_o)$. Définir de la même manière comme positifs en y les $n_{01} = n(x_0, y_1)$. Ainsi, $(n_{10} - n_{01})/n = (p_x - p_y)$ représentera la différence nette dans la proportion des éléments positifs en x et y. Les éléments qui sont positifs ou négatifs à la fois en x et y ne comptent pas dans les différences. Nous obtenons ainsi un cas de trois catégories comme en (1) et en (2). On peut citer à titre d'exemple la différence entre les proportions de gens qui "interdiraient tous les essais nucléaires" et ceux qui "souhaitent un désarmement nucléaire complet", ou entre ceux qui "obligeraient les Irakiens à quitter le Koweit" et ceux qui voudraient "remplacer Saddam" (Wild et Seber 1993). Toutefois, les deux catégories pourraient également être tirées de deux *sondages différents* portant sur les mêmes n cas, si on avait affaire par exemple à une vérification de la qualité, à des observations provenant de deux bases de sondage ou aux deux vagues d'un même échantillon. Ces situations ressemblent à celles examinées aux points (4) et (5).

Comparaisons du type B

4. Les expressions test-retest et avant-après qualifient des plans dans lesquels les mêmes sujets font l'objet de deux observations. Les réponses dichotomiques $n_2 = n_{10}$ désignent alors le nombre de changements négatifs; $n_0 = n_{01}$ le nombre de changements positifs et $n_{11} + n_{00}$ la somme des positifs et des négatifs inchangés. On dénote respectivement les réponses positives et négatives par 1 et 0, et les première et deuxième vagues par l'ordre de l'indice. La différence $(n_{10} + n_{11}) - (n_{01} + n_{11}) = n_{10} - n_{01} = n_2 - n_0$ représente une mesure du changement entre les positifs pour les deux observations, et $p_2 - p_0 = n_2/n - n_0/n$ représente la mesure du changement des proportions (McNemar 1949; Cochran 1950; Mosteller 1952).

5. Les paires appariées de n paires de sujets peuvent également être traitées comme une généralisation du plan test-retest (Mosteller 1952). Par exemple, n paires de sujets randomisés peuvent représenter le traitement expérimental par rapport au traitement témoin; ou n paires de garçons ou de filles par rapport aux variables de contrôle. Le traitement statistique $(p_{10} - p_{01})$ des n paires de sujets appariés est le même que celui des n paires de traitements du même groupe de n sujets (4).

La similitude des traitements statistiques de ces cinq plans de deux types distincts est pratique; nous présentons les résultats empiriques obtenus pour les deux types. "Elle a également une valeur heuristique qu'on a ignorée dans des publications récentes (Scott et Seber 1983; Wild et Seber 1993). Les résultats des tests de chi carré pour les types 4 et 5 ont été publiés depuis longtemps (McNemar 1949; Cochran 1950; Mosteller 1952) et la similitude des cas des catégories 1, 2 et 3 a également été démontrée" (Kish 1965, 12.10). (Kish faisait erreur lorsqu'il a parlé de "trinômes et de binômes appariés" pour désigner les "trichotomies et les dichotomies appariées" puisque ces termes s'appliquent uniquement aux échantillons IID.)

Tous ces traitements portent sur les différences des proportions p_i fondées sur les variables de dénombrement n_i. Les extensions aux différences corrélées $(y_i - y_j)$ pour d'autres variables sont envisageables, mais elles sortent du cadre de la présente étude. On pourrait citer à titre d'exemple pratique la différence entre les parts en dollars (pas seulement les nombres n_i) entre deux marques d'automobiles sur un total de $\sum y_i$ ventes.

3. ERREURS D'ÉCHANTILLONNAGE ET EFFETS DU PLAN DE SONDAGE

Pour les échantillons aléatoires simples de taille n, on peut facilement montrer (Kish 1965, 12.10) que

$$\mathrm{var}(p_2 - p_0) = \left[\frac{(1-f)n}{(n-1)}\right]\left[p_2 + p_0 - (p_2 - p_0)^2\right]/n.$$

La plupart des exemples trouvés et utilisés sont tirés de grands échantillons d'enquêtes où la valeur de $(1 - f)$ peut être ignorée. Il convient de noter que pour la variance des éléments

$$p_2 + p_0 - (p_2 - p_0)^2 = p_2q_2 + p_0q_0 + 2p_2p_0,$$

où le dernier terme $\mathrm{cov}(p_2,p_0) = -p_2p_0$ désigne la covariance découlant de ce que p_2 et p_0 sont des parties concurrentes du même échantillon, plutôt que les proportions d'échantillons indépendants. Le carré de la différence de proportions $(p_2 - p_0)^2$ sera habituellement un terme de correction petit; en l'excluant, nous obtiendrons l'équivalent de la variance $(p_2 + p_0)/n$ de deux échantillons de Poisson indépendants. On se rappellera en outre (Kish 1965, 12.10) que:

> Le test de chi carré a été utilisé pour certains de ces problèmes traités séparément (Cochran 1950; Mosteller 1952; McNemar 1962, p. 225). Il s'agit essentiellement de $(n_2 - n_0)^2/(n_2 + n_0)$, c'est-à-dire du carré de la différence divisé par la variance, en vertu de l'hypothèse nulle $n_2 = n_0$. On applique exactement les théories disponibles pour les tests d'hypothèses nulles dans les petits échantillons, y compris la "correction de Yates", toutes fondées sur l'hypothèse d'un échantillonnage aléatoire simple. Toutefois, le traitement de ces problèmes dans de grands échantillons avec des moyennes estimatives et des écarts-types adéquats présente d'énormes avantages. Premièrement, au lieu de nous limiter à tester les hypothèses nulles, nous pouvons faire des déductions fondées sur les intervalles de probabilité $(p_2 - p_0) \pm t_p$ *é.-t.* $(p_2 - p_0)$. Deuxièmement, les formules des écarts-types d'échantillons complexes peuvent s'appliquer directement à la moyenne $(p_2 - p_0)$. Troisièmement, la structure logique de cette statistique $(p_2 - p_0)$ s'observe plus facilement dans son application à plusieurs problèmes distincts.

Les proportions corrélées proviennent habituellement de données tirées d'enquêtes complexes, et les calculs de la variance devraient être appropriés pour le plan de sondage. On peut adopter les formules de la variance propres aux échantillons complexes stratifiés, mais la formule directe comporte huit termes (Kish 1965, 12.10.3). Au lieu de cela, il est plus pratique de traduire le problème en une variable trichotomique avec des valeurs de -1, 0, $+1$, comme dans le plan 2 de la section 2. Les calculs de la section 4 utilisent cette transformation.

Les comparaisons entre les variables et entre les échantillons peuvent alors être facilitées par le recours aux effets du plan de sondage:

$$\mathrm{eps}^2(p_2 - p_0) = \frac{\text{variance calculée de } (p_2 - p_0)}{[p_2 + p_0 - (p_2 - p_0)^2]/n}.$$

Il convient de s'attarder quelque peu sur les limites de l'utilisation des eps en guise d'outils pour les approximations robustes. Ils sont utiles pour les échantillons en grappes et les échantillons à plusieurs degrés qui utilisent les sélections primaires pour le calcul des erreurs d'échantillonnage. Toutefois, nous avons voulu éviter le problème des échantillons pondérés parce que leur traitement serait trop spécifique et peut-être trop complexe. La pondération pour l'absence de réponse ne serait pas importante pour le rapport de $\mathrm{eps}(p_i - p_j)$ sur $\mathrm{eps}(p_i)$. Toutefois, la pondération des inégalités majeures des probabilités de sélection appelle des traitements particuliers. Néanmoins, la déduction et l'expérience portent à conclure que les valeurs d'eps sont moins sensibles aux poids que les variances et les moyennes elles-mêmes. Par ailleurs, nous présumons que les rapports que nous avons observés entre les valeurs d'$\mathrm{eps}(p_i - p_j)$ et d'$\mathrm{eps}(p_i)$ se maintiendront avec les données pondérées, à condition que ces dernières ne soient pas extrêmes ni pathologiques.

Il serait utile de déterminer un rapport approximatif, mais fiable, d'$\mathrm{eps}(p_i - p_j)$ sur $\mathrm{eps}(p_i)$ et $\mathrm{eps}(p_j)$ afin de pouvoir déduire la valeur du premier, peu aisée à calculer, à partir des valeurs des seconds que l'on peut obtenir plus facilement. Plusieurs conjectures de rechange peuvent paraître raisonnables à cet égard; aucune ne peut être dérivée mathématiquement ni exclue.

1. $\mathrm{Eps}(p_i - p_j) = 1$, l'absence d'effet du plan de sondage, a été implicitement présumée dans les cinq publications citées à la section 1.

2. $\mathrm{Eps}(p_i) > \mathrm{eps}(p_i - p_j) > 1$, désignant des effets persistants, mais moins importants que les $\mathrm{eps}(p_i)$ pour les proportions. C'est ce qui se produit dans le cas des "classes croisées" et de leurs comparaisons (Kish 1987, 7.1). Cette hypothèse a également paru raisonnable à plusieurs statisticiens d'expérience que nous avons consultés.

3. $\mathrm{Eps}(p_i - p_j) = [\mathrm{eps}(p_i) + \mathrm{eps}(p_j)]/2$ nous a semblé être une bonne approximation pour l'ensemble de nos données pour diverses populations et divers plans de sondage. Cette hypothèse nous paraît raisonnable puisque les effets du plan de sondage dus au groupement des valeurs individuelles de p_i peuvent s'appliquer de façon similaire à la variable issue de la différence $(p_i - p_j)$ entre deux de ces valeurs.

4. Des résultats incohérents seraient également possibles, mais peu utiles puisqu'ils empêcheraient toute déduction.

4. RÉSULTATS EMPIRIQUES POUR EPS$(P_i - P_j)$

À défaut d'un solide fondement théorique ou mathématique permettant d'accorder une préférence à l'une ou l'autre des conjectures énumérées plus haut, les résultats empiriques sur les valeurs d'eps$(p_i - p_j)$ deviennent essentiels puisqu'ils permettent d'établir un rapport entre ces valeurs et les valeurs calculées d'eps(p_i). Ceci rappelle nos hypothèses plus familières concernant eps$(p_i) = \sqrt{1 + roh[\bar{b} - 1]}$; leur valeur dépend de plusieurs facteurs influant sur *roh*, le coefficient de corrélation intraclasse, ainsi que de la taille moyenne de la grappe $\bar{b}$ (Kish 1965, 5.4, 8.2). Les valeurs d'eps(p_i) varient beaucoup d'un sondage à l'autre, ainsi que d'une variable à l'autre dans un sondage particulier (Kish, Groves et Krotki 1976; Verma, Scott et O'Muircheartaigh 1980; Verma et Lê 1995). Toutefois, les études empiriques portant sur les erreurs d'échantillonnage observées dans diverses enquêtes constituent, pour les statisticiens d'enquête, une source utile de connaissances qui leur permettent également d'établir des rapports entre les valeurs d'eps de statistiques complexes et celles d'eps(p_i) plus simples (Kish 1995; Rao et Wu 1985; Rao et Scott 1987). De même, pour en savoir plus sur le rapport entre eps$(p_i - p_j)$ et eps(p_i), nous pouvons utiliser les résultats empiriques issus de plusieurs variables et de plusieurs enquêtes.

Dans un premier essai portant sur cette question, nous présentons les données tirées de quatorze enquêtes réalisées dans une vaste gamme de situations. Onze de ces enquêtes présentent cinq ensembles de données (figures 1 et 2 et tableaux 1 à 3) et portent sur les différences de catégories appariées tirées d'enquêtes uniques (type A). Trois ensembles de résultats (tableaux 1 à 3) proviennent d'enquêtes sociales, et les deux autres (figures 1 et 2) proviennent d'enquêtes sur la démographie et la santé des populations. Finalement, trois autres ensembles comportant chacun deux vagues de données sont fondés sur deux nouvelles séries d'entrevues menées auprès des mêmes répondants (tableaux 4, 5 et 6), et représentent les plans de comparaison du type B.

Tableaux:

1. La National Election Study de 1986, réalisée par le Institute for Social Research de la University of Michigan, $n = 2{,}135$.
2. La National Education Longitudinal Study (NELS) de 1988, réalisée par le National Opinion Research Center de la University of Chicago, $n = 24{,}355$.
3. La National Longitudinal Study of Labor Market Experience of Youth, réalisée par le National Opinion Research Center de la University of Chicago, $n = 5{,}857$.
4. Les National Election Studies Panels de 1990 et de 1992, réalisés par le Survey Research Center, Institute for Social Research, Ann Arbor, Michigan 48106.
5. Les Panel Study of Income Dynamics de 1983 et de 1987, réalisés par le Survey Research Center.
6. Les études Americans' Changing Lives de 1986 et de 1989 réalisées par le Survey Research Center.

Figures:

1. Les études sur la démographie et la santé réalisées au Maroc, au Niger et en Colombie par MACRO International.
2. La portion du recensement indonésien portant sur la strate rurale de Java (données inédites).

Nous relevons en particulier dans ces études les résultats EMPIRIQUES suivants:

1) D'abord et avant tout: les effets du plan de sondage sur les différences eps$(p_i - p_j)$ ne sont habituellement PAS MOINS IMPORTANTS que les eps(p_i) sur les proportions elles-mêmes, et eps$(p_i - p_j) \simeq 0.5$ [eps(p_i) + eps(p_j)] dans tous les cas. Ils varient de concert avec l'importante variation des valeurs d'eps entre les variables, et avec les variations moins importantes observées entre les paires de catégories d'une même variable. Les chercheurs qui négligent les eps font l'erreur commune de sous-estimer les erreurs d'échantillonnage propres aux statistiques d'enquêtes par grappes. Outre l'intérêt qu'elle présente, cette observation fournit un modèle utile de déduction puisque les trois autres sources de variation - entre les variables, entre les catégories de chaque variable et erreurs d'échantillonnage des statistiques individuelles - sont toutes plus importantes.

2) On peut obtenir ces résultats avec les 14 ensembles de données d'enquêtes des tableaux et des graphiques, et en donner une illustration au tableau 1. Noter que les eps varient essentiellement entre 1.00 pour la variable D (problèmes dans le pays) et 2.32 pour la variable A (religion), ce qui nous donne une valeur d'eps^2 = $2.32^2 = 5.38$. Il est rassurant de constater que notre règle empirique (1) se vérifie pour toute la gamme des valeurs. On observe fréquemment une telle variation entre les variables d'un même échantillon; ceci devrait nous convaincre de cesser d'utiliser une moyenne commune pour tous les eps d'un échantillon (Verma et Lê 1995; Kish 1995).

   Il convient par ailleurs de souligner la grande variation des valeurs d'eps pour les cinq catégories de la même variable, soit de 1.21 à 2.32 (n° 3 pour les protestants "intégristes"). Il en découle que la valeur d'eps$(p_i - p_j)$ n'est grande que lorsque *i* ou *j* appartiennent à la catégorie 3 pour cette variable. Ces variations des eps entre les catégories de la même variable signifient qu'il vaudrait mieux les calculer pour toutes les catégories plutôt que pour une seule catégorie "représentative". La possibilité d'une grande variation entre les catégories pour une seule et même variable est une conclusion importante de notre étude qui était semble-t-il passée inaperçue jusqu'à maintenant.

3) Des erreurs d'échantillonnage sont également présentes dans le calcul des eps. Il semble que seuls les statisticiens qui ont une bonne expérience du traitement des erreurs d'échantillonnage et des effets du plan de sondage puissent avoir une idée de l'importance de ces erreurs. Ces erreurs risquent d'être les principales responsables

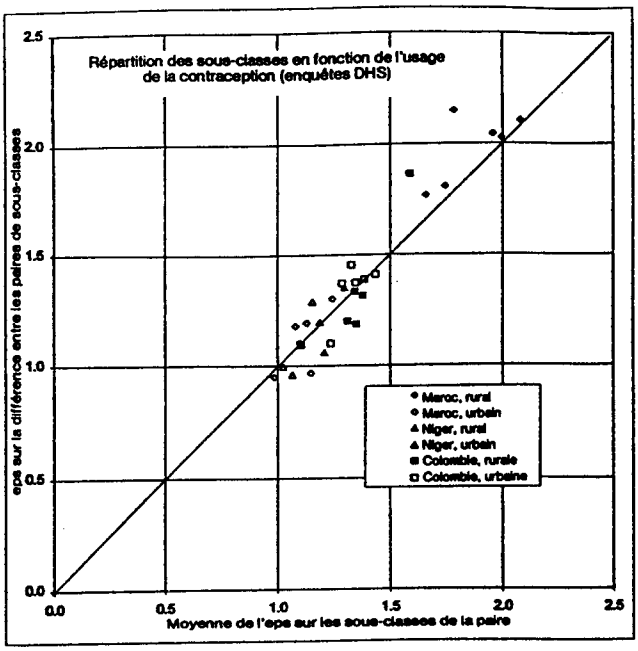

Figure 1. Comparaison de eps$(p_i - p_j)$ à la moyenne de eps(p_i) et de eps(p_j) pour diverses catégories d'usage de la contraception*. Illustration de six populations tirées des enquêtes sur la démographie et la santé des populations (DHS).

* 1 = aucune méthode de contraception
2 = usage des méthodes traditionnelles uniquement
3 = utilisation d'une méthode moderne "réversible"
4 = stérilisation.

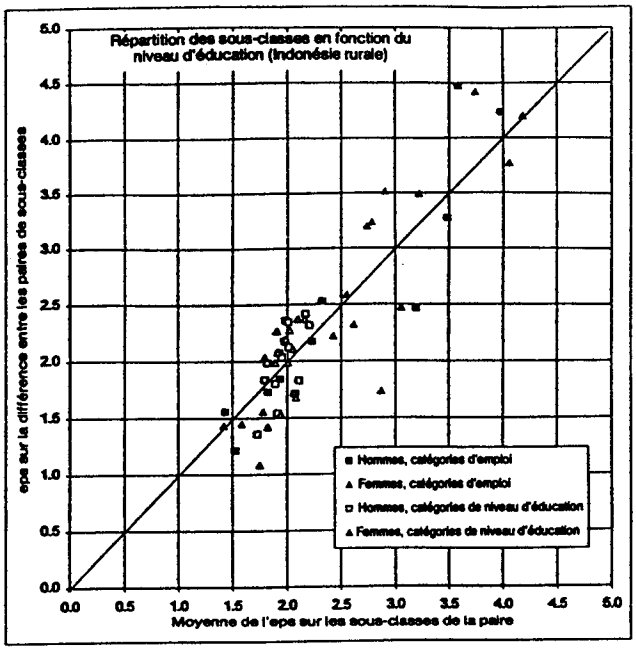

Figure 2. Comparaison de eps$(p_i - p_j)$ à la moyenne de eps(p_i) et de eps(p_j) pour les diverses catégories d'emploi et de niveau d'éducation par sexe. Illustration d'un recensement.

Tableau 1

La National Election Study de 1986 réalisée par le Institute for Social Research de la University of Michigan (n = 2,135)

Catégories	Eps pour			
$i - j$	P_i	P_j	Moyenne	$(P_i - P_j)$
A. Religion				
1-2	1.21	1.42	1.32	1.10
1-3	1.21	2.32	1.77	2.02
1-4	1.21	1.50	1.36	1.18
1-5	1.21	1.18	1.19	1.17
2-3	1.42	2.32	1.87	1.93
2-4	1.42	1.50	1.46	1.57
2-5	1.42	1.18	1.30	1.27
3-4	2.32	1.50	1.91	2.03
3-5	2.32	1.18	1.75	2.04
4-5	1.50	1.18	1.34	1.19
Moyenne	*1.56*	*1.53*	*1.54*	*1.55*
C. Appui à Reagan				
1-2	1.32	1.10	1.21	1.07
1-3	1.32	.86	1.09	1.26
1-4	1.32	1.48	1.40	1.50
2-3	1.10	.86	.98	.96
2-4	1.10	1.48	1.29	1.38
3-4	.86	1.48	1.17	1.09
Moyenne	*1.17*	*1.21*	*1.19*	*1.21*
Moyenne globale	*1.23*	*1.24*	*1.23*	*1.24*

Catégories	Eps pour			
$i - j$	P_i	P_j	Moyenne	$(P_i - P_j)$
B. Position sur l'avortement				
1-2	1.27	.97	1.12	.97
1-3	1.27	1.28	1.28	1.32
1-4	1.27	1.31	1.29	1.36
2-3	.97	1.28	1.12	1.08
2-4	.97	1.31	1.14	1.16
3-4	1.28	1.31	1.30	1.32
Moyenne	*1.17*	*1.24*	*1.21*	*1.20*
D. Problèmes dans le pays				
1-2	1.07	.94	1.00	.98
1-3	1.07	1.04	1.05	1.09
1-4	1.07	.93	1.00	1.12
2-3	.94	1.04	.99	1.01
2-4	.94	.93	.93	.85
3-4	1.04	.93	.98	.82
Moyenne	*1.02*	*.97*	*.99*	*.98*

Tableau 2

La National Education Longitudinal Study (NELS) de 1988, réalisée par le National Opinion Research Center de la University of Chicago (n = 24,355)

Catégories	Eps pour			
$i - j$	P_i	P_j	Moyenne	$(P_i - P_j)$
A. Scolarité				
1-2	1.38	1.22	1.30	1.11
1-3	1.38	1.14	1.26	1.16
1-4	1.38	1.19	1.29	1.30
1-5	1.38	1.42	1.40	1.54
2-3	1.22	1.14	1.18	1.11
2-4	1.22	1.19	1.21	1.24
2-5	1.22	1.42	1.32	1.45
3-4	1.14	1.19	1.17	1.18
3-5	1.14	1.42	1.28	1.37
4-5	1.19	1.42	1.31	1.20
Moyenne	*1.27*	*1.28*	*1.27*	*1.25*
E. Religion				
1-2	2.48	2.83	2.65	2.74
1-3	2.48	2.02	2.25	2.09
2-3	2.83	2.02	2.42	2.59
Moyenne	*2.60*	*2.29*	*2.44*	*2.47*
I. Assurance				
1-2	1.42	1.28	1.35	1.37
B. Trouve les cours ennuyants				
1-2	.99	1.11	1.05	1.04
1-3	.99	1.12	1.06	1.07
2-3	1.11	1.12	1.12	1.13
Moyenne	*1.03*	*1.12*	*1.08*	*1.08*
C. Libre de choisir sa voie				
1-2	1.28	1.10	1.19	1.21
1-3	1.28	1.08	1.18	1.28
2-3	1.10	1.08	1.09	.97
Moyenne	*1.22*	*1.09*	*1.15*	*1.15*
D. L'école donne accès à de bons emplois				
1-2	1.24	1.07	1.16	1.17
1-3	1.24	1.11	1.18	1.24
2-3	1.07	1.11	1.09	1.01
Moyenne	*1.18*	*1.10*	*1.14*	*1.14*
F. Scolarité du père				
1-2	1.61	1.76	1.69	1.83
1-3	1.61	1.68	1.65	1.65
2-3	1.76	1.68	1.72	2.48
Moyenne	*1.65*	*1.71*	*1.69*	*1.99*
Moyenne globale	*1.48*	*1.41*	*1.45*	*1.49*

Tableau 3

La National Longitudinal Study of Labor Market Experience of Youth, réalisée par le National Opinion Research Center de la University of Chicago (n = 5,857)

Catégories	Eps pour			
$i - j$	P_i	P_j	Moyenne	$(P_i - P_j)$
A. La chance joue un rôle important dans ma vie				
1-2	1.26	1.20	1.23	1.06
1-3	1.26	1.18	1.22	1.30
1-4	1.26	1.16	1.21	1.28
2-3	1.20	1.18	1.19	1.22
2-4	1.20	1.16	1.18	1.25
3-4	1.18	1.16	1.17	1.05
Moyenne	*1.23*	*1.17*	*1.20*	*1.19*
C. Je fais ce que je veux dans la vie				
1-2	1.13	1.06	1.10	1.09
1-3	1.13	1.10	1.12	1.13
2-3	1.06	1.10	1.08	1.07
Moyenne	*1.11*	*1.09*	*1.10*	*1.10*
E. Mes projets aboutissent rarement				
1-2	1.19	1.07	1.13	1.12
1-3	1.19	1.13	1.16	1.20
2-3	1.07	1.13	1.10	1.08
Moyenne	*1.15*	*1.11*	*1.13*	*1.13*
I. Travail de la mère				
1-2	1.49	1.36	1.43	1.41
1-3	1.49	1.52	1.51	1.53
2-3	1.36	1.52	1.44	1.44
Moyenne	*1.45*	*1.47*	*1.46*	*1.47*
Moyenne globale	*1.20*	*1.17*	*1.18*	*1.19*
B. Quelque chose m'empêche de progresser				
1-2	1.07	1.22	1.14	1.04
1-3	1.07	1.12	1.10	1.14
1-4	1.07	1.09	1.08	1.09
2-3	1.22	1.12	1.17	1.28
2-4	1.22	1.09	1.16	1.14
3-4	1.12	1.09	1.11	1.07
Moyenne	*1.13*	*1.12*	*1.13*	*1.13*
D. Je suis aussi capable que les autres				
1-2	1.17	1.13	1.15	1.16
1-3	1.17	1.07	1.12	1.16
2-3	1.13	1.07	1.10	1.08
Moyenne	*1.16*	*1.09*	*1.12*	*1.13*
F. Je suis satisfait				
1-2	1.19	1.12	1.16	1.16
1-3	1.19	1.13	1.16	1.20
2-3	1.12	1.13	1.13	1.09
Moyenne	*1.17*	*1.13*	*1.15*	*1.15*

des quelques cas où eps$(p_i - p_j)$ ne se situe pas entre eps(p_i) et eps(p_j), et où on obtient eps$(p_i) <$ eps$(p_i - p_j) >$ eps(p_j) ou eps$(p_i) >$ eps$(p_i - p_j) <$ eps(p_j). Ces cas montrent en passant que nos résultats ne sont pas des conséquences mathématiques, mais qu'ils ont plutôt un caractère empirique.

Tableau 4

Les National Election Studies Panels de 1990 et de 1992, réalisés par le Survey Research Center, Institute for Social Research, Ann Arbor

Catégories avant/après (90/92)	Eps pour			
	P_i	P_j	Moyenne	$(P_i - P_j)$
Entièrement d'accord avec Bush	1.14	.93	1.04	1.02
D'accord avec la politique étrangère de Bush	.92	1.05	.99	1.00
Désapprouve entièrement la politique étrangère de Bush	1.23	1.24	1.24	1.32
D'accord avec la politique économique de Bush	.97	.94	.96	.96
Entièrement d'accord avec la politique économique de Bush	1.14	1.04	1.09	1.10
D'accord avec Bush	1.00	1.00	1.00	1.00
Désapprouve entièrement Bush	1.16	1.10	1.13	1.12
A suivi la campagne électorale à la télé	.89	1.55	1.22	1.40
Moyenne	*1.06*	*1.11*	*1.08*	*1.11*

Tableau 5

Les Panel Study of Income Dynamics de 1983 et de 1987, réalisés par le Survey Research Center, Ann Arbor

Catégories* avant/après (83/87)	Eps pour			
	P_i	P_j	Moyenne	$(P_i - P_j)$
Vit dans le sud	1.22	1.23	1.23	1.11
Âge du chef de famille	1.28	1.33	1.31	1.37
Taille de la famille	1.29	1.43	1.36	1.47
Nombre d'enfants dans la famille	1.23	1.43	1.33	1.49
Heures de travail du chef de famille	1.12	.84	.98	1.03
Âge de l'enfant le plus jeune	.93	.91	.92	.87
Moyenne	*1.18*	*1.20*	*1.19*	*1.22*

* Toutes les variables sont classées dans deux catégories.

Les résultats empiriques présentés aux figures 1 et 2 viennent encore confirmer ceux présentés aux tableaux 1, 2 et 3. Nous y constatons également que: 1) eps$(p_i - p_j) \cong$ [eps(p_i) + eps(p_j)]/2 approximativement, le long de la ligne de 45°; 2) ces égalités se vérifient pour une vaste gamme d'eps; 3) la variation entre les variables est en effet très importante. Cette variation est particulièrement évidente pour l'Indonésie rurale, où les valeurs d'eps dépassent 4 et où les valeurs d'eps^2 sont donc supérieures à 16. Ces importants effets de groupement sont dus à la grande taille des grappes: avec $\bar{b} = 133$ et 137, les valeurs de $roh = 0.12$ sont suffisantes pour un grand eps. À noter que ces résultats empiriques viennent à la fois de populations et de variables très diversifiées; elles sont différentes l'une de l'autre et différentes également des données des

Tableau 6

Les études Americans' Changing Lives de 1986 et de 1989 réalisées par le Survey Research Center, Ann Arbor

Catégories avant/après	Eps pour			
	P_i	P_j	Moyenne	$(P_i - P_j)$
A. Rencontre des amis				
Une fois par semaine	1.30	1.26	1.28	1.28
2 ou 3 fois par mois	.88	1.00	.94	1.02
Moyenne	*1.09*	*1.13*	*1.11*	*1.15*
B. Fait de l'exercice physique				
Souvent	1.51	1.67	1.59	1.26
Jamais	1.62	1.97	1.80	1.41
Moyenne	*1.56*	*1.82*	*1.70*	*1.34*
C. Satisfaction de soi				
Très satisfait	1.28	1.21	1.25	1.33
Pas satisfait	1.04	1.16	1.10	1.00
Moyenne	*1.16*	*1.19*	*1.18*	*1.17*
D. Aime sa maison				
Beaucoup	1.24	.90	1.07	.91
Pas beaucoup	1.33	.98	1.16	1.12
Moyenne	*1.29*	*.94*	*1.12*	*1.02*
E. Pratique le jardinage				
Souvent	1.40	1.16	1.28	1.19
Rarement	.91	1.11	1.01	1.18
Jamais	1.66	1.17	1.42	1.26
Moyenne	*1.32*	*1.15*	*1.24*	*1.21*
F. A une attitude positive				
Oui	1.10	1.33	1.22	1.19
Non	1.05	1.28	1.17	1.21
Moyenne	*1.08*	*1.31*	*1.20*	*1.20*
Moyenne globale	*1.25*	*1.26*	*1.26*	*1.18*

tableaux 1, 2 et 3. La figure 1 présente les données provenant de trois pays (Maroc, Niger et Colombie); il y a donc six populations puisque les eps en zone urbaine sont assez différents de ceux observés en zone rurale. La figure 2 présente les résultats selon le sexe; on distingue deux populations passablement distinctes en ce qui concerne l'emploi, mais la différence est moins marquée en ce qui concerne le niveau d'éducation.

Nous attendions les données empiriques des tableaux des études 4, 5 et 6 avec anxiété. Il est vrai que les cinq ensembles précédents avaient conduit à des conclusions similaires, même s'ils traitaient de onze populations différentes, de pointages et de variables. Pourtant, les études 1 à 5 portaient sur des paires de catégories issues de polytomies, les plans 1 et 2 du type A. Dorénavant, nous cherchions des données pour des comparaisons de type B provenant d'enquêtes par panel, de manière à pouvoir explorer les conjectures des plans test-retest et avant-après. Du point de vue mathématique, on peut facilement montrer une ressemblance avec les polytomies (c.-à-d., avec les tétratomies), mais les rapports entre ces valeurs et les valeurs empiriques des eps ne sont pas du tout évidentes. C'est ce qui explique pourquoi ces valeurs empiriques sont si utiles et si remarquables. Nous avons ici observé un effet du plan de sondage extrêmement important pour les tests de chi carré des comparaisons analytiques.

5. SIGNIFICATION DES RÉSULTATS POUR LES RECHERCHES CONNEXES

Nos travaux antérieurs et ceux d'autres chercheurs fournissent déjà une masse considérable de données empiriques sur les effets du plan de sondage sur l'échantillon total, les sous-classes et les différences, pour des variables et des plans différents. Il serait donc utile d'examiner les résultats décrits plus haut à la lumière de ces données antérieures.

On a déterminé que la nature des variables d'enquête qui font l'objet d'une estimation est un facteur important (souvent le plus important) pour déterminer l'ampleur des eps. On peut observer des eps extrêmement différents selon le type de variables, même au sein d'un seul échantillon ou avec des plans de sondage très semblables. C'est la raison pour laquelle nous avons toujours recommandé que les eps soient calculés pour plusieurs variables différentes, alors qu'il est généralement moins important de les calculer pour plusieurs sous-classes différentes, en particulier lorsque ces sous-classes sont définies en fonction des mêmes caractéristiques.

Nos résultats montrent que les eps peuvent également varier énormément entre les catégories différentes de la même variable d'enquête, lorsque l'échantillon total sert de base commune à l'estimation. Ainsi, il convient dans un certain sens de considérer chaque catégorie individuelle et chaque différence entre les paires de catégories, même lorsqu'on a affaire aux variables d'une seule et même enquête, comme s'il s'agissait de variables séparées aux fins du calcul et de l'analyse des effets du plan de sondage.

Pour ce qui est du rapport existant entre les eps pour les sous-classes et pour les différences entre sous-classes, les recherches antérieures ont surtout porté sur la situation décrite ci-après. Compte tenu d'un échantillon total n réparti en sous-classes i de taille $n_i = p_i.n$, les valeurs d'eps(r_i) pour les statistiques r_i (comme le rapport m_i/n_i, la moyenne $\sum y_i/n_i$ ou le rapport $\sum y_i/\sum x_i$), calculées sur la base des éléments de sous-classe n_i, se rapportent à la valeur eps(r) pour la même variable, calculée sur la base de l'échantillon total. De même, les valeurs de eps$(r_i - r_j)$ pour les différences de sous-classes se rapportent à l'eps(r_i), à l'eps(r_j) fondé sur les sous-classes individuelles et à l'eps(r) fondé sur l'échantillon total. De nombreux calculs confirment que ces rapports s'accordent avec notre conjecture (2) de la section 3:

$$\text{eps}(r) > \text{eps}(r_i); \quad \text{et} \quad \text{eps}(r_i) > \text{eps}(r_i - r_j) > 1.$$

Ces effets des covariances sur les effets du plan de sondage pour des échantillons par grappes sont essentiellement empiriques (même sociologiques, au sens large); ils doivent être vérifiés comme tels.

Ainsi en est-il du nouveau rapport que nous avons découvert pour $(p_i - p_j)$ pour deux catégories, lesquelles sont si différentes de ce qui précède. Les rapports eps$(p_i - p_j) \cong [\text{eps}(p_i) + \text{eps}(p_j)]/2$ sont également empiriques et approximatifs, et doivent être constamment vérifiés. Pourtant, ils semblent s'appliquer largement à nos données et sont nettement préférables aux autres hypothèses, telles que eps$(p_i - p_j) = 1$, qu'on a souvent utilisées jusqu'ici.

REMERCIEMENTS

Les auteurs désirent remercier l'éditeur associé et les arbitres dont les commentaires ont permis d'abréger et d'améliorer cet article.

BIBLIOGRAPHIE

COCHRAN, W.G. (1950). The comparison of percentages in matched samples. *Biometrika*, 37, 256-66.

DEMING, W.E. (1953). On the distinction between enumerative and analytic studies. *Journal of the American Statistical Association*, 48, 244-45.

KISH, L. (1965). *Survey Sampling*. New York: John Wiley and Sons.

KISH, L. (1987). *Statistical Research Design*. New York: John Wiley and Sons.

KISH, L. (1995). Methods for design effects. *Journal of Official Statistics*, 11, 55-77.

KISH, L., et FRANKEL, M.R. (1974). Inference from complex samples. *Journal of the Royal Statistical Society*, (B), 36, 1-37.

KISH, L., GROVES, R.M., et KROTKI, K. (1976). *Sampling Errors for Fertility Surveys*. Document hors série n° 17, *Enquête mondiale de la fécondité*. Institut International de Statistique: The Hague.

LÊ, T., et VERMA, V. (1995). *Sample Designs and Sampling Errors for the DHS*. Calverton MD: MACRO International.

McNEMAR, Q. (1949). *Psychological Statistics*. New York: John Wiley and Sons.

MOSTELLER, F. (1952). Some statistical problems in measuring the subjective responses to drugs. *Biometrika*, 8, 220-226.

RAO, J.N.K., et SCOTT, A.J. (1987). On simple adjustments to Chi-square tests with sample survey data. *Annals of Statistics*, 15, 385-397.

RAO, J.N.K., et WU, C.F.S. (1985). Inference from stratified samples: Second-order analysis of three methods for nonlinear statistics. *Journal of the American Statistical Association*, 80, 620-630.

SCOTT, A.J., et HOLT, D. (1982) The effect of two-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association*, 77, 848-54.

SCOTT, A.J., et SEBER, G.A.F. (1983). Difference of proportions from the same survey. *The American Statistician*, 37, 319-20.

SKINNER, C.J., HOLT, D., et SMITH, T.M.F. (1989). *Analysis of Complex Surveys*. New York: John Wiley and Sons.

VERMA, V., et LÊ, T. (1995). Sampling errors for the DHS survey. 50ième Session de l'Institut International de Statistique, Beijing.

VERMA, V., SCOTT, C., et O'MUIRCHEARTAIGH, C. (1980). Sample designs and sampling errors for the World Fertility Surveys. *Journal of the Royal Statistical Society (A)*, 143, 431-473.

WILD, C.J., et SEBER, G.A.F. (1993). Comparing two proportions for the same survey. *The American Statistician*, 47, 178-181.