Techniques d'enquête, décembre 1998
Vol. 24, n° 2, pp. 209-218
Statistique Canada

Intervalles de confiance pour les proportions à petit nombre d'événements positifs prévus estimées au moyen des données d'enquête

EDWARD L. KORN et BARRY I. GRAUBARD[1]

RÉSUMÉ

En dehors des enquêtes complexes, on applique fréquemment des intervalles de confiance «exacts» aux proportions obtenues par distribution binomiale au lieu d'intervalles reposant sur une normalité approximative, quand les événements positifs sont peu nombreux. Malheureusement, il est impossible d'en faire autant avec les données des enquêtes complexes, en sorte qu'on utilise des intervalles reposant sur la normalité approximative hypothétique de la proportion pondérée selon l'échantillon, même quand il existe peu d'événements positifs. Les auteurs proposent une simple modification aux intervalles binomiaux dans une situation de ce genre. Les simulations restreintes qu'ils présentent indiquent que la couverture probable de tels intervalles dépasse celle des intervalles reposant sur la normalité, des intervalles venant de la transformation par logit et des intervalles résultant de l'approximation de Poisson. Ils appliquent ensuite ces intervalles à la prévalence du virus de l'immunodéficience humaine (VIH) selon les données de la troisième Enquête nationale sur la santé et la nutrition et à la proportion de cocaïnomanes selon les données de l'Enquête sur la santé et la nutrition des personnes hispaniques.

MOTS CLÉS: Intervalle de confiance binomial; intervalle de confiance exact; transformation par logit; intervalle confiance de Poisson.

1. INTRODUCTION

Avec les données d'une enquête complexe, l'intervalle de confiance de niveau $1 - \alpha$ typique d'une proportion d'événements positifs pour une variable de type 0-1 ressemble à

$$\hat{p} \pm t_d\,(1 - \alpha/2)\,[\hat{\text{var}}(\hat{p})]^{1/2} \qquad (1.1)$$

où $\hat{p}$ représente l'estimateur de la proportion pondéré pour l'échantillon, où $\hat{\text{var}}(\hat{p})$ est l'estimateur de la variance de $\hat{p}$, et où $t_d(1 - \alpha/2)$ correspond au $1 - \alpha/2$ quantile d'une distribution t à d degrés de liberté. On calcule l'estimateur $\hat{\text{var}}(\hat{p})$ par linéarisation ou par répétition au moyen d'une méthode tenant compte du plan d'échantillonnage, notamment le fait que $\hat{p}$ est un estimateur pondéré. Par données d'une enquête complexe, on entend les données recueillies en vertu d'un plan d'échantillonnage à plusieurs degrés, avec sélection stratifiée de grappes au premier degré. Avec un plan d'échantillonnage de ce genre, d correspond habituellement au nombre de grappes échantillonnées moins le nombre de strates (Korn et Graubard 1990). L'intervalle de confiance (1.1), que nous appellerons «intervalle linéaire» repose sur l'hypothèse que $\hat{p}$ a une distribution proche de la normale. L'hypothèse se vérifie pour diverses asymptotes raisonnables (Krewski et Rao 1981). L'emploi du quantile t au lieu d'un quantile à distribution normale dans (1.1) s'appuie sur des preuves empiriques (Frankel 1971, ch. 7), mais il est possible de le justifier formellement à partir d'hypothèses fermes (Korn et Graubard 1990).

Avec un nombre restreint d'événements positifs cependant, $\hat{p}$ perd sa normalité approximative (Cochran 1977, p. 58). On peut éviter l'hypothèse de la normalité en utilisant l'intervalle de confiance de Clopper et Pearson (1934) reposant sur la distribution binomiale lorsqu'on est en présence d'un échantillon aléatoire simple (ou qu'il ne s'agit pas d'une enquête complexe); Vollset (1993) donne une analyse complète des intervalles de confiance des proportions dans un tel contexte. Quand un échantillon aléatoire simple de taille n comporte x valeurs positives, l'intervalle de confiance Clopper et Pearson $(p_L(x, n), p_U(x, n))$ de niveau $1 - \alpha$ peut s'exprimer comme suit (Johnson, Kotz et Kemp 1993, p. 130):

$$p_L(x, n) = \frac{v_1 F_{v_1, v_2}(\alpha/2)}{v_2 + v_1 F_{v_1, v_2}(\alpha/2)}$$

$$p_U(x, n) = \frac{v_3 F_{v_3, v_4}(1 - \alpha/2)}{v_4 + v_3 F_{v_3, v_4}(1 - \alpha/2)} \qquad (1.2)$$

où $v_1 = 2x$, $v_2 = 2\,(n - x + 1)$, $v_3 = 2(x + 1)$, $v_4 = 2\,(n - x)$ et $F_{d_1, d_2}(\beta)$ correspond au quantile β d'une distribution F à d_1 et d_2 degrés de liberté. Pour un intervalle de confiance unilatéral, on utilise α au lieu de $\alpha/2$ dans les expressions ci-dessus. On sait qu'avec un échantillon aléatoire simple, la probabilité de couverture de tels intervalles est supérieure ou égale à leur niveau nominal, peu importe le nombre d'événements positifs. On parle parfois d'intervalles de confiance «exacts», mais nous dirons plutôt «intervalles binomiaux».

[1] Edward L. Korn, Biometric Research Branch, EPN-739, National Cancer Institute, Bethesda, MD 20892, U.S.A.; Barry I. Graubard, Biostatistics Branch, National Cancer Institute, Bethesda, MD 20892, U.S.A.

Nous proposons une simple modification aux intervalles binomiaux susceptible d'en permettre l'application aux proportions estimées à partir des données d'une enquête complexe. Nous nous intéressons particulièrement au cas où un petit nombre d'événements positifs est prévu. Beaucoup d'analystes d'enquête ne fourniront pas de proportions estimées en pareil cas, car elles manqueraient de fiabilité. Si on appliquait le critère de l'erreur-type relative aux proportions de l'Enquête-ménages nationale sur la toxicomanie de 1996 (SAMHSA 1998), par exemple, la proportion estimée de femmes cocaïnomanes n'apparaîtrait pas au tableau 7. Or, nous pensons que de telles proportions véhiculent une information intéressante pourvu qu'on en précise de façon explicite l'imprécision en donnant les intervalles de confiance. On trouvera à la partie 2 les intervalles que nous proposons et ceux reposant sur la transformation par logit et la distribution de Poisson suggérés par d'autres auteurs. À la partie 3, les résultats d'une simulation comparent l'efficacité des intervalles en question. On constate que les intervalles proposés donnent de bons résultats pour ce qui est de la probabilité de couverture de la proportion réelle et de la largeur moyenne de l'intervalle. La partie 4 présente deux applications des intervalles dans une enquête majeure où l'on s'attend à un faible nombre d'événements positifs. L'article se termine par une analyse de travaux connexes débouchant sur des intervalles de confiance qui atteindront leur probabilité de couverture nominale quelle que soit la configuration des groupes au sein de la population.

2. LIMITES DES INTERVALLES PROPOSÉS ET DES AUTRES INTERVALLES DE CONFIANCE

Pour l'intervalle de confiance $1-\alpha$ d'un échantillon de taille n, définissons d'abord la taille efficace de l'échantillon par

$$n^* = \frac{\hat{p}(1-\hat{p})}{\text{vâr}(\hat{p})} \tag{2.1}$$

et la taille efficace de l'échantillon ajustée selon le nombre de degrés de liberté par

$$n^*_{df} = \frac{\hat{p}(1-\hat{p})}{\text{vâr}(\hat{p})}\left(\frac{t_{n-1}(1-\alpha/2)}{t_d(1-\alpha/2)}\right)^2 \tag{2.2}$$

Les valeurs n^* et n^*_{df} sont égales à n quand $\hat{p}=0$. Les limites proposées remplacent n par n^*_{df}, et x par $\hat{p}n^*_{df}$ dans (1.2), *viz.* $p_L(\hat{p}n^*_{df}, n^*_{df})$ et $p_U(\hat{p}n^*_{df}, n^*_{df})$. (Quand n est élevé, on peut se servir du quantile $1-\alpha/2$ de la distribution normale au lieu de $t_{n-1}(1-\alpha/2)$ dans (2.2).) Pour estimer l'intervalle de confiance d'une proportion dans un sous-groupe de la population, on suppose que la taille de l'échantillon n est égale à celle de l'échantillon restreint au sous-groupe.

Voici la justification heuristique de la procédure qui précède. La taille efficace de l'échantillon (2.1) est n divisée par un estimateur de l'effet du plan de sondage de l'enquête. La chose paraît raisonnable si on veut tenir compte de la variabilité supplémentaire de $\hat{p}$ attribuable à l'échantillonnage complexe. La variabilité de l'estimateur de la variance présente elle aussi de l'importance lorsqu'on construit les intervalles de confiance. La deuxième fraction de (2.2) tient compte du fait que $\text{vâr}(\hat{p})$ variera typiquement plus qu'un estimateur de la variance utilisé avec un échantillonnage aléatoire simple. Si d est élevé, ce facteur prend presque la valeur un et est inutile. Avec une faible valeur d et une valeur n et $\hat{p}n^*_{df}$ élevées, il serait préférable que l'intervalle proposé se rapproche de l'intervalle (1.1), qui convient en pareil cas. Puisque $F_{u,w}(\beta) \cong 1 + z(\beta)\sqrt{2(1/u + 1/w)}$ pour les valeurs u et w élevées (Johnson et Kotz 1970, p. 81), c'est effectivement ce qui se produit, à savoir $\hat{p} - p_L(\hat{p}n^*_{df}, n^*_{df}) \cong p_U(\hat{p}n^*_{df}, n^*_{df}) - \hat{p} \cong t_d(1-\alpha/2)[\text{vâr}(\hat{p})]^{1/2}$.

Breeze (1990) a élaboré une méthode analogue à celle que nous suggérons pour l'Enquête générale sur les ménages du Royaume-Uni. Cette méthode repose sur l'intervalle de confiance $1-\alpha$ de l'échantillonnage aléatoire simple $(po_L(x), po_U(x))$ pour une variable de Poisson x, s'exprimant comme suit (Johnson et coll. 1993, p. 171):

$$po_L(x) = 0.5\,\chi^2_{v_1}(\alpha/2) \text{ et } po_U(x) = 0.5\,\chi^2_{v_2}(1-\alpha/2)$$

où $v_1 = 2x$, $v_2 = 2(x+1)$ et $\chi^2_v(\beta)$ correspond au quantile β d'une distribution χ^2 à v degrés de liberté. On suppose que l'intervalle de confiance est égal à $(po_L(\hat{p}n^*)/n^*, po_U(\hat{p}n^*)/n^*)$ avec les données d'une enquête complexe.

Une troisième façon de bâtir un intervalle de confiance consiste à recourir à la transformation par logit. L'intervalle de confiance de niveau $1-\alpha$ correspond à

$$\left(\frac{1}{1+\exp(-LLOGIT)}, \frac{1}{1+\exp(-ULOGIT)}\right)$$

où

$$LLOGIT = \log\frac{\hat{p}}{1-\hat{p}} - t_d(1-\alpha/2)\frac{[\text{vâr}(\hat{p})]^{1/2}}{\hat{p}(1-\hat{p})} \tag{2.3}$$

et

$$ULOGIT = \log\frac{\hat{p}}{1-\hat{p}} + t_d(1-\alpha/2)\frac{[\text{vâr}(\hat{p})]^{1/2}}{\hat{p}(1-\hat{p})} \tag{2.4}$$

On a suggéré d'utiliser ces intervalles avec un quantile à distribution normale plutôt qu'à distribution t pour l'Enquête-ménages nationale sur la toxicomanie de 1996 (SAMHSA 1998). En dehors des conditions d'enquête complexe, quand $\hat{p}=0$, on pourrait ajouter une petite constante au nombre d'événements et de non-événements observés, par ex. 1/2, afin de calculer l'intervalle de

confiance de la transformation par logit (Agresti 1990, pp. 249-250). Dans le cas présent, quand $\hat{p} = 0$, l'intervalle de confiance a été fixé à l'intervalle binomial $(p_L(0, n), p_U(0, n))$.

Lorsqu'on sait que l'effet (véritable) du plan d'échantillonnage dépassera 1 avant l'échantillonnage, il est possible d'apporter diverses modifications aux procédures qui précèdent. Nous préconisons pour notre part de tronquer à n la taille efficace de l'échantillon ajustée en fonction du nombre de degrés de liberté. Bref, si n^*_{df} est plus grand que n, nous en fixons la valeur à n et prenons comme limites supérieure et inférieure de l'intervalle $p_L(\hat{p}n, n)$ et $p_U(\hat{p}n, n)$. Pour les intervalles de Breeze, on pourrait établir que n^* est égal à n si $n^* > n$. Avec l'intervalle linéaire ou à logit, il est possible de prendre l'estimateur de la variance de l'échantillon aléatoire simple $\hat{p}(1-\hat{p})/n$ au lieu de $\hat{\text{var}}(\hat{p})$ dans (1.1), (2.3) et (2.4) si $n^* > n$; on trouvera d'autres troncations dans SAMHSA (1998). On justifie ces troncations par le fait que l'instabilité de l'estimateur de la variance $\hat{\text{var}}(\hat{p})$ peut donner un effet de plan estimé inférieur à un. L'instabilité en question peut s'avérer particulièrement importante puisque $\hat{p}$ est petit (SAMHSA 1998). Ces méthodes de troncation ont pour conséquence d'élargir les intervalles de confiance et de les rendre plus conservateurs. En théorie, on pourrait aussi ajuster la taille réelle estimatée de l'échantillon si on sait que l'effet (véritable) du plan de sondage est inférieur à un avant l'échantillonnage. Par mesure de prudence, nous recommandons toutefois de l'éviter.

Dans cet article, nous nous concentrons sur les intervalles de confiance de la probabilité de «superpopulation» que $Y = 1$ et n'égale pas la proportion de la population finie. En d'autres termes, le paramètre visé est $p = \sum_{u=1}^{N} p_u / N$ et non pas $P = \sum_{u=1}^{N} Y_u / N$, où Y_u présente une distribution de Bernoulli de paramètre p_u, et N correspond à la taille de la population. Les probabilités de couverture obtenues par simulation qui apparaissent à la section suivante se rapportent donc à la couverture de p. Ce paramètre étant établi, nous ne recourrons pas aux facteurs de correction de la population finie au moment d'estimer $\hat{\text{var}}(\hat{p})$ utilisé dans (2.2); nous ne procéderons pas ici à d'autres ajustements de la variance $\hat{\text{var}}(\hat{p})$ attribuable au plan d'échantillonnage pour l'inférence à la superpopulation (Korn et Graubard 1998). Un examinateur a suggéré qu'on recoure à un modèle pour estimer l'intervalle de confiance de p. Bien que limités, nos essais indiquent néanmoins qu'une approche de ce genre donne des estimateurs semblables aux estimateurs pondérés et ne présente aucun avantage sur le plan de l'inférence (Pfeffermann et LaVange 1989; Graubard et Korn 1996).

Ceux qui s'intéressent à un intervalle de confiance pour P devraient se servir des intervalles proposés en intégrant des facteurs de correction pour la population finie à $\hat{\text{var}}(\hat{p})$, dans (2.2). On pourrait établir un intervalle de confiance pour $\sum_{u=1}^{N} Y_u$ en multipliant les extrémités de l'intervalle de P par N, s'il est connu, ou par l'estimateur $\hat{N}$ de N, dans le cas contraire. (Théoriquement, on pourrait aussi tenir compte de la variabilité de $\hat{N}$, mais la variabilité supplémentaire sera faible.) Une autre façon d'estimer l'intervalle de confiance de P consisterait à modifier les limites usuelles (Guenther 1983) d'un échantillon aléatoire simple (d'après la distribution hypergéométrique), un peu comme les intervalles suggérés modifient les intervalles binomiaux.

3. SIMULATIONS

Les principaux résultats de la simulation apparaissent aux tableaux 1 à 5. Le tableau 1 donne les résultats des simulations pour 32 grappes de 100 unités chacune. Dans la grappe i, on a simulé le nombre d'événements positifs au moyen d'une distribution binomiale assortie de p_i comme paramètre probabiliste. Les valeurs $\{p_i, i = 1, ..., 32\}$ du tableau 1 correspondent à la probabilité dans les grappes. Dans le tiers supérieur du tableau, cette probabilité est égale à la constante p = 0,1, 0,02, 0,01 ou 0,0025, ce qui correspond à un nombre d'événements positifs prévu de 320, 64, 32 ou 8 sur un total possible de 3 200. Dans le tiers intermédiaire du tableau, la probabilité associée aux grappes est égale à $p/2$ avec une probabilité de 1/2 ou à $3p/2$ avec une probabilité de 1/2, p ayant la même valeur que dans le premier tiers. En faisant varier p_i d'une grappe à l'autre, une corrélation se développe entre les observations de la grappe. Cette corrélation est égale à 0,00278, 0,0051, 0,0025 et 0,0006 dans le tiers intermédiaire (quand on néglige les poids de l'échantillonnage) et correspond à 320, 64, 32 ou 8 événements positifs prévus, respectivement. Pour le tiers inférieur du tableau, la probabilité est égale à $p/2$ avec une probabilité de 3/4 ou à $5p/2$, avec une probabilité de 1/4, ce qui correspond à une corrélation intraclasse de 0,0833, 0,0153, 0,0076 et 0,0019. Un poids d'échantillonnage de 1 ou de 10 est attribué au hasard dans les simulations du tableau 1, avec une probabilité de 1/2 pour chaque observation (poids sans valeur informative).

Les résultats du tableau 1 conviennent aux limites inférieure et supérieure d'un intervalle de confiance unilatéral de 95 %; idéalement, le pourcentage représentant la non-couverture dans le tableau devrait être inférieur ou égal à la valeur nominale de 5,0. Les résultats conviennent aussi aux intervalles bilatéraux de 90 %, pour lesquels, idéalement, les valeurs supérieure et inférieure du tableau devraient être ≤5,0 (Jennings 1987). Pour chaque ligne du tableau, on a simulé 100 000 jeux de données au moyen du générateur de nombre aléatoire du SAS (1990, p. 631), de manière à estimer la probabilité de non-couverture des limites de l'intervalle.

Tableau 1
Simulation de la non-couverture (en pour cent) des limites supérieure et inférieure d'un intervalle de confiance unilatéral de 95 % comptant 32 grappes de 100 observations chacune; poids d'échantillonnage de 1 ou de 10 avec probabilité de 1/2 (poids non informatifs)

Distribution des proportions pour les grappes[a]	Proportion globale	Nombre prévu positif	Méthode de calcul des limites de l'intervalle							
			Linéaire		Logit		Breeze		Proposée	
			Infér.	Sup.	Infér.	Sup.	Infér.	Sup.	Infér.	Sup.
(1)										
0,1	0,1	320	4,6	5,5	5,3	4,6	4,5	4,1	4,8	4,4
0,02	0,02	64	3,4	7,1	5,2	4,6	4,5	4,7	4,2	4,4
0,01	0,01	32	2,9	8,0	5,4	4,5	4,4	4,5	4,0	4,1
0,0025	0,0025	8	1,6	9,5	5,5	1,8	3,6	2,2	3,3	1,8
(1/2, 1/2)										
0,05, 0,15	0,1	320	4,3	5,8	5,5	4,3	4,3	3,8	4,7	4,1
0,01, 0,03	0,02	64	3,1	7,5	5,2	4,8	4,3	4,8	4,0	4,5
0,005, 0,015	0,01	32	2,7	8,6	5,2	4,7	4,1	4,9	3,7	4,4
0,00125, 0,00375	0,0025	8	1,5	9,9	5,4	2,0	3,4	2,3	3,1	2,0
(3/4, 1/4)										
0,05, 0,25	0,1	320	3,1	7,8	4,7	5,6	3,4	5,0	3,6	5,3
0,01, 0,05	0,02	64	2,7	8,6	5,1	5,3	4,0	5,4	3,7	5,0
0,005, 0,025	0,01	32	2,2	9,8	5,0	5,3	3,7	5,5	3,3	5,0
0,00125, 0,00625	0,0025	8	1,3	10,7	5,3	2,2	3,3	2,5	3,0	2,2

(a) Les fractions entre parenthèses indiquent la probabilité que la proportion de grappes corresponde à la valeur mentionnée.

Tableau 2
Simulation de la non-couverture (en pour cent) des limites supérieure et inférieure d'un intervalle de confiance de 95 % pour un échantillon de 32 grappes de 100 observations chacune; les poids d'échantillonnage sont informatifs et correspondent à 1 ou 10 (voir le corps du texte)

Distribution des proportions pour les grappes[a]	Proportion pondérée globale	Nombre prévu positif	Méthode de calcul des limites de l'intervalle							
			Linéaire		Logit		Breeze		Proposée	
			Infér.	Sup.	Infér.	Sup.	Infér.	Sup.	Infér.	Sup.
(1)										
0,1	0,1	191,0	4,3	5,9	5,1	4,9	4,2	4,4	4,6	4,6
0,02	0,02	36,9	3,3	7,3	5,3	4,3	4,4	4,4	4,1	4,1
0,01	0,01	18,4	2,8	8,7	5,5	4,0	4,3	4,3	3,9	3,7
0,0025	0,0025	4,6	1,3	18,7	6,1	4,8	3,2	4,8	2,8	4,8
(1/2, 1/2)										
0,05, 0,15	0,1	191,0	5,0	5,0	6,4	3,7	5,1	3,2	5,4	3,4
0,01, 0,03	0,02	36,9	3,0	7,9	5,4	4,5	4,3	4,6	4,0	4,3
0,005, 0,015	0,01	18,4	2,5	9,2	5,4	4,2	4,1	4,4	3,7	3,9
0,00125, 0,00375	0,0025	4,6	1,3	19,0	6,1	4,9	3,2	4,9	2,8	4,9
(3/4, 1/4)										
0,05, 0,25	0,1	191,0	4,7	5,7	7,1	4,1	5,1	3,6	5,5	3,8
0,01, 0,05	0,02	36,9	2,6	8,9	5,2	5,2	4,0	5,3	3,7	4,9
0,005, 0,025	0,01	18,4	2,1	10,1	5,3	4,8	3,8	5,1	3,4	4,5
0,00125, 0,00625	0,0025	4,6	1,2	19,8	5,9	5,3	3,2	5,3	2,8	5,3

(a) Les fractions entre parenthèses indiquent la probabilité que la proportion de grappes pondérée corresponde à la valeur mentionnée.

Tableau 3
Simulation de la non-couverture (en pour cent) des limites supérieure et inférieure d'un intervalle de confiance unilatéral de 95 % pour un échantillon de 32 grappes de 100 observations chacune; analyses non pondérées

Distribution des proportions pour les grappes[a]	Proportion globale	Nombre prévu positif	Méthode de calcul des limites de l'intervalle							
			Linéaire		Logit		Breeze		Proposée	
			Infér.	Sup.	Infér.	Sup.	Infér.	Sup.	Infér.	Sup.
(1)										
0,1	0,1	320	5,0	4,9	5,7	4,2	4,9	3,8	5,2	4,1
0,02	0,02	64	3,8	6,3	5,2	4,5	4,7	4,8	4,4	4,4
0,01	0,01	32	3,5	6,8	5,6	4,4	4,7	4,4	4,3	4,0
0,0025	0,0025	8	2,5	8,8	5,6	3,8	4,1	3,9	3,9	3,9
(1/2, 1/2)										
0,05, 0,15	0,1	320	4,5	5,6	5,6	4,2	4,5	3,7	4,8	4,0
0,01, 0,03	0,02	64	3,4	7,0	5,1	4,8	4,5	4,9	4,1	4,6
0,005, 0,015	0,01	32	3,0	7,6	5,2	4,8	4,4	4,8	3,9	4,4
0,00125, 0,00375	0,0025	8	2,2	9,2	5,4	4,3	3,8	4,3	3,5	4,3
(3/4, 1/4)										
0,05, 0,25	0,1	320	3,3	7,7	4,8	5,6	3,5	5,1	3,7	5,3
0,01, 0,05	0,02	64	2,9	8,1	5,1	5,2	4,1	5,3	3,8	4,9
0,005, 0,025	0,01	32	2,5	9,2	4,9	5,6	3,9	5,6	3,5	5,2
0,00125, 0,00625	0,0025	8	2,0	10,4	5,3	5,1	3,8	5,1	3,3	5,1

(a) Les fractions entre parenthèses indiquent la probabilité que la proportion de grappes corresponde à la valeur mentionnée.

Tableau 4
Simulation de la non-couverture (en pour cent) des limites supérieure et inférieure d'un intervalle de confiance unilatéral de 95 % pour un échantillon de 32 grappes de 10 observations chacune; poids d'échantillonnage de 1 ou 10 avec probabilité de 1/2 (poids non informatifs)

Distribution des proportions pour les grappes[a]	Proportion globale	Nombre prévu positif	Méthode de calcul des limites de l'intervalle							
			Linéaire		Logit		Breeze		Proposée	
			Infér.	Sup.	Infér.	Sup.	Infér.	Sup.	Infér.	Sup.
(1)										
0,2	0,2	64	4,0	6,6	5,2	4,7	3,1	4,7	4,2	4,3
0,1	0,1	32	3,2	7,8	5,3	4,4	3,6	3,8	3,9	4,0
0,025	0,025	8	1,7	10,2	5,5	2,1	3,4	2,1	3,2	2,4
(1/2, 1/2)										
0,1, 0,3	0,2	64	3,6	7,0	5,0	4,9	2,8	3,4	3,9	4,4
0,05, 0,15	0,1	32	3,0	8,1	5,1	4,6	3,4	4,0	3,7	4,2
0,0125, 0,0375	0,025	8	1,6	10,6	5,4	2,1	3,3	2,1	3,1	2,5
(3/4, 1/4)										
0,1, 0,5	0,2	64	3,1	7,8	4,6	5,3	2,4	3,9	3,3	4,8
0,05, 0,25	0,1	32	2,5	9,2	4,8	5,2	3,0	4,6	3,3	4,8
0,0125, 0,0625	0,025	8	1,5	11,5	5,3	2,4	3,2	3,5	3,0	2,8

(a) Les fractions entre parenthèses indiquent la probabilité que la proportion de grappes corresponde à la valeur mentionnée.

Tableau 5

Simulation de la non-couverture (en pour cent) des limites inférieure et supérieure d'un intervalle de confiance unilatéral de 95 % pour un échantillon de 32 grappes de 10 ou de 100 observations chacune avec probabilité de 1/2; poids d'échantillonnage de 1 ou 10 avec probabilité de 1/2 (poids non informatifs)

Distribution des proportions pour les grappes[a]	Proportion globale	Nombre prévu positif	Méthode de calcul des limites de l'intervalle							
			Linéaire		Logit		Breeze		Proposée	
			Infér.	Sup.	Infér.	Sup.	Infér.	Sup.	Infér.	Sup.
(1)										
0,1818	0,1818	320	5,1	6,0	5,7	5,2	4,2	4,1	5,2	5,0
0,0364	0,0364	64	4,1	7,6	5,7	5,2	5,0	5,2	4,8	4,9
0,0182	0,0182	32	3,4	8,5	5,7	5,0	4,7	5,1	4,4	4,7
0,0045	0,0045	8	2,0	12,7	5,9	3,4	4,0	4,3	3,6	3,8
(1/2, 1/2)										
0,0909, 0,2727	0,1818	320	5,0	6,4	6,1	4,8	4,2	3,6	5,2	4,4
0,0182, 0,0545	0,0364	64	3,9	8,1	6,0	5,1	4,9	5,0	4,7	4,8
0,0091, 0,0273	0,0182	32	3,1	9,3	5,8	5,2	4,5	5,3	4,2	4,9
0,0023, 0,0068	0,0045	8	1,8	13,2	5,9	3,6	3,9	4,5	3,5	4,0
(3/4, 1/4)										
0,0909, 0,4545	0,1818	320	3,1	9,9	4,6	7,6	2,5	6,3	3,3	7,1
0,0182, 0,0909	0,0364	64	2,8	10,9	5,3	7,3	3,9	7,3	3,7	7,0
0,0091, 0,0455	0,0182	32	2,4	11,5	5,4	6,8	3,9	6,9	3,6	6,5
0,0023, 0,0114	0,0045	8	1,6	14,5	5,7	4,0	3,7	5,0	3,3	4,4

(a) Les fractions entre parenthèses indiquent la probabilité que la proportion de grappes pondérée corresponde à la valeur mentionnée.

La limite supérieure de l'intervalle de confiance linéaire tombe plus en-dessous de sa valeur véritable que le niveau nominal de 5 %, même avec un compte prévu de jusqu'à 320 événements positifs, en particulier quand il y a une corrélation positive dans la grappe (tiers intermédiaire et inférieur du tableau), ce qui est quelque peu étonnant. Dans le cas de l'intervalle de confiance de la transformation par logit, la non-couverture semble dépasser légèrement le niveau nominal, surtout pour les limites inférieures. Les limites des intervalles de Breeze et des intervalles proposés paraissent généralement conservatrices. Celles des intervalles binomiaux avec échantillonnage aléatoire simple ne conviennent pas aux cas simulés au tableau 1, à cause des poids d'échantillonnage et de la corrélation dans la grappe (dans les deux tiers inférieurs du tableau). Une non-couverture de plus de 8 % pour les limites inférieure et supérieure de l'intervalle binomial pour tous les cas du tableau, en est la preuve (résultats non indiqués).

Discuter de la «longueur» des intervalles de confiance unilatéraux s'avère un peu compliqué aussi nous restreindrons-nous à parler de la longueur des intervalles bilatéraux à 90 %. Pour l'ensemble des simulations du tableau 1, les intervalles de Breeze et ceux que nous proposons ont une largeur moyenne de 3,3 % et de 4,9 % supérieure à celle des intervalles de transformation par logit.

Le tableau 2 présente les résultats d'une simulation semblable à celle du tableau 1, si ce n'est que les poids de l'échantillon sont informatifs. Pour cela, on a fixé le poids à 10 avec probabilité de 2/3 si l'événement était positif et de 1/3 dans le cas contraire, sinon le poids était de 1. La probabilité que chaque grappe comprenne un événement positif a été ajustée à la baisse jusqu'à ce que les proportions pondérées générales soient identiques à celles du tableau 1. Les résultats du tableau 2 ressemblent à ceux du tableau 1 mais les intervalles linéaire et à logit ont tendance à donner de moins bonnes probabilités de couverture.

Au tableau 3, on peut voir les résultats d'une simulation identique à celle du tableau 1, à l'exception du fait que l'analyse n'est pas pondérée. Comme le tiers supérieur du tableau 3 n'indique aucune corrélation entre les éléments des grappes, on pourrait aussi bien recourir aux limites binomiales de l'échantillon aléatoire simple. Si on effectue la moyenne des quatre possibilités dans ce tiers du tableau, on constate que les limites proposées donnent une largeur de 2,5 % supérieure à celle de l'intervalle binomial (résultats non indiqués). Puisque l'effet véritable du plan de sondage dans le tiers supérieur du tableau 3 est égal à 1,0, les simulations permettent d'examiner les conséquences de la troncation de n_{df}^{*} dans la méthode proposée. (Il y a rarement troncation dans les simulations du tableau 1, car l'incidence véritable du plan de sondage est toujours >1.) Lorsqu'on applique la méthode suggérée avec troncation, la simulation donne des intervalles plus larges et plus conservateurs que ceux du premier tiers du tableau 3. Après calcul de la moyenne des quatre cas, on se

rend compte que les limites proposées créent un intervalle de 4,0 % plus large avec troncation, comparativement à sans (résultats non indiqués pour les limites tronquées).

Le tableau 4 montre les résultats d'une simulation identique à celle du tableau 1, sauf pour le nombre d'observations par grappe, qui passe de 100 à 10. Les résultats ressemblent fort à ceux du tableau 1 pour les simulations comprenant le même nombre prévu d'événements positifs. Seuls les intervalles de Breeze, plus conservateurs que ceux de la méthode préconisée, font exception à la règle. On le doit au fait que la proportion globale est plus importante au tableau 4 qu'au tableau 1, pour le même nombre d'événements prévus (l'échantillon du tableau 4 étant plus petit). Les intervalles de Breeze modifiés par l'approximation de Poisson ne donnent pas de bons résultats avec les grandes proportions. Pour le vérifier, nous avons effectué une simulation identique à celle du tiers supérieur du tableau 1 mais en prenant p = 0,5 comme proportion pour 1 600 événements positifs prévus. Les pourcentages inférieur et supérieur de non-couverture était de 1,2 % et 1,3 %, respectivement, contre 4,6 % et 4,7 % avec la méthode proposée. En moyenne, les intervalles de Breeze sont 37 % plus larges que les intervalles proposés.

Les intervalles de Breeze laissent aussi à désirer avec un très petit nombre de grappes, car ils négligent le nombre de degrés de liberté dans l'estimation de la variance. Nous avons, par exemple, effectué une simulation correspondant au premier tiers du tableau 1 en ne prenant les données que de 8 grappes (100 observations par grappe), et $p_i \equiv 0{,}1$, de telle sorte que le nombre d'événements positifs prévus se situait à 80. Les pourcentages inférieur et supérieur de non-couverture s'établissaient à 6,1 % et à 5,4 %, respectivement, pour les intervalles de Breeze, comparativement à 4,7 % et à 4,0 % pour la méthode préconisée.

Le tableau 5 donne les résultats d'une simulation identique à celle du tableau 1, mais avec des grappes de 10 ou de 100, selon une probabilité de 1/2. Les probabilités de non-couverture dépassent la valeur nominale de 5 % du tiers inférieur du tableau, peu importe la méthode employée. En outre, les intervalles à logit ne se comportent pas aussi bien qu'au tableau 1 dans les deux premiers tiers du tableau.

On a procédé à une dernière simulation reproduisant deux grappes (de 50) pour chaque strate. Le nombre d'événements positifs prévu était identique à celui du tableau 1, les poids ont été fixés au hasard à 1 ou à 10 et la probabilité d'un événement positif variait avec les strates, cela afin d'imiter une corrélation à l'intérieur de la grappe. Les résultats (non indiqués) se rapprochaient beaucoup de ceux du tableau 1.

4. APPLICATIONS

Nous examinerons maintenant deux applications caractérisées par un petit nombre d'événements positifs. La première concerne l'estimation de l'infection par le VIH au sein d'une population non sélectionnée. Le nombre de réagissants est peu élevé, car le taux d'infection par le VIH est relativement faible. Dans le deuxième cas (estimation de la cocaïnomanie), les taux sont élevés mais le nombre d'événements positifs reste faible parce que l'analyse ne s'applique qu'à de petits sous-groupes. Dans les deux cas, on a utilisé SUDAAN (Shah, Barnwell et Bieler 1995) pour calculer l'erreur-type (attribuable au plan de sondage) des proportions et la fonction «FINV» du SAS (1990, p. 547) pour établir les quantiles de la distribution F en (1.2).

4.1 Estimation de la séroprévalence du VIH dans le cadre de la troisième Enquête nationale sur la santé et la nutrition (NHANES III)

L'enquête NHANES III s'est déroulée de 1988 à 1994 et couvrait les membres de la population civile non intitutionnalisés de 2 mois et plus aux États-Unis (National Center for Health Statistics 1994). Les participants de 18 ans et plus ont subi un test de dépistage du VIH. McQuillan, Khare, Karon, Schable et Vlahov (1997) se sont penchés sur la prévalence du virus chez les personnes de moins de 60 ans et dans divers sous-groupes, dont certains sont reproduits au tableau 6. Sur les 11 202 personnes testées, 59 portaient le virus. La prévalence estimée au tableau 6 (0,32 %) s'écarte considérablement de la proportion non pondérée (0,53 % = 59/11 202) parce qu'elle correspond à une proportion pondérée par les poids de l'échantillon. Les tests de dépistage préservant l'anonymat, les poids d'échantillonnage ont été dérivés des poids originaux de l'échantillon de la NHANES III pour les personnes testées présentant les mêmes caractéristiques, soit mêmes îlot (lieu de l'enquête), race ou groupe ethnique, sexe et groupe d'âge (18-39 et 40-59 ans) (communication personnelle de M. Khare). Le pseudo-plan de sondage utilisé pour estimer la variance consistait à prélever 2 pseudo-UPE dans chacune des 23 strates (communication personnelle de M. Khare), ce qui ne correspond pas au pseudo-plan de sondage typique servant à estimer la variance dans la NHANES III.

Les intervalles de confiance linéaires de 90 % s'appliquant à la prévalence du VIH dans les différents groupes, au tableau 6, sont décalés à gauche et sont plus étroits que les autres intervalles, qui se ressemblent. Les intervalles proposés sont légèrement plus larges que ceux de Breeze ou à logit. La taille efficace de l'échantillon établie en (2.1) est nettement plus faible que celle de l'échantillon, à cause des effets du plan de sondage; les intervalles de confiance obtenus après troncation seront donc identiques à ceux du tableau 6. L'écart entre n^* et n_{df}^* s'avère relativement faible. Pour cet événement positif relativement rare, les simulations décrites à la partie 3 donnent à penser que les intervalles de confiance de Breeze et ceux proposés maintiennent mieux la probabilité de couverture nominale de 90 % que les autres intervalles.

Tableau 6
Séroprévalence du VIH chez les adultes de 18 à 59 ans selon la troisième Enquête nationale sur la santé et la nutrition

	Total	Sexe		Race/ascendance		
		Masculin	Féminin	Blanc	Noir	Mex.-Amér.
Taille de l'échantillon	11202	5142	6060	4128	3579	3495
Nombre infectés	59	44	15	9	38	12
Prévalence (%) ± ET	0,320 ± 0,076	0,519 ± 0,130	0,127 ± 0,053	0,203 ± 0,071	1,100 ± 0,247	0,368 ± 0,134
Taille efficace de l'éch.						
n^*	5588	3056	4433	3976	1779	2039
n^*_{df}	5148	2816	4084	3664	1640	1880
Int. conf. linéaire 90 %	(0,19, 0,45)	(0,30, 0,74)	(0,04, 0,22)	(0,08, 0,33)	(0,68, 1,52)	(0,14, 0,60)
Int. conf. logit 90 %	(0,21, 0,48)	(0,34, 0,80)	(0,06, 0,26)	(0,11, 0,37)	(0,75, 1,62)	(0,20, 0,69)
Int. conf. Breeze 90 %	(0,21, 0,48)	(0,32, 0,79)	(0,05, 0,26)	(0,10, 0,37)	(0,73, 1,61)	(0,18, 0,68)
Int. conf. proposé 90 %	(0,20, 0,48)	(0,32, 0,80)	(0,05, 0,26)	(0,10, 0,37)	(0,71, 1,63)	(0,17, 0,69)

Tableau 7
Adultes de 12 à 44 ans qui ont essayé la cocaïne d'après la population d'hispanophones de 16 ans et plus de scolarité échantillonnée dans le cadre de l'Enquête sur la santé et la nutrition des personnes d'origine hispanique

	Total	Sexe	
		Masculin	Féminin
Taille de l'échantillon	123	69	54
A essayé	13	10	3
Proportion (%) ± ET	11,6 ± 2,5	14,3 ± 3,4	7,0 ± 4,8
Taille efficace de l'éch.			
n^*	1671	1050	282
n^*_{df}	1328	844	229
Int. conf. linéaire 90 %	(7,0, 16,2)	(8,0, 20,7)	(-1,9[a], 15,9)
Int. conf. logit 90 %	(7,8, 17,1)	(9,1, 21,9)	(1,9, 22,8)
Int. conf. Breeze 90 %	(7,7, 17,0)	(8,3, 23,2)	(0,9, 24,8)
Int. conf. proposé 90 %	(7,4, 17,2)	(8,5, 22,1)	(0,9, 22,7)
	Avec troncation		
Int. conf. linéaire 90 %	(6,3, 17,0)	(6,5, 22,2)	voir ci-dessus
Int. conf. logit 90 %	(7,2, 18,2)	(8,1, 24,1)	"
Int. conf. Breeze 90 %	(7,1, 18,1)	(7,7, 24,4)	"
Int. conf. proposé 90 %	(7,2, 17,5)	(8,0, 23,2)	"

(a) En pratique, ces intervalles auraient pour limites (0, 15,9) puisque les proportions ne peuvent être négatives.

4.2 Cocaïnomanie chez les diplômés de collège échantillonnés lors de l'Enquête sur la santé et la nutrition des personnes d'origine hispanique (HHANES)

La HHANES s'est déroulée en 1982-1983 et portait sur trois groupes hispanophones des États-Unis (National Center for Health Statistics 1985). Nous ne nous sommes intéressés qu'à l'échantillon Mexicain-Américain. On a demandé aux personnes de 12 à 44 ans à quel âge elles avaient essayé de la cocaïne pour la première fois. Les réponses possibles étaient l'âge du répondant lorsqu'il avait essayé de la cocaine pour la première fois (en années), «n'a jamais essayé» et «ne sait pas». Nous avons tenté d'estimer la proportion de ceux qui avait déjà fait l'expérience de la drogue parmi les personnes comptant 16 années ou plus de scolarité (pour lesquelles la réponse «ne sait pas» n'existait pas).

Sur les 123 personnes échantillonnées, 13 avait déjà absorbé de la cocaïne, soit une proportion de 11,6 % après pondération (tableau 7). L'erreur-type attribuable au plan de sondage (2,5 %) n'est estimée qu'avec 8 degrés de liberté parce qu'on peut reproduire approximativement le plan de sondage de l'enquête en prélevant 2 UPE dans chacune des 8 strates (Kovar et Johnson 1986). La taille efficace de l'échantillon s'établit à $n^* = 167{,}1$ et $n^*_{df} = 132{,}8$, soit plus que l'échantillon, dans les deux cas. On le doit à l'effet estimé du plan de sondage (0,73, de sorte que $n^* = 123/0{,}736 = 167{,}1$). (Le deuxième facteur de (2.2) est égal à 0,794.) Malgré la stratification, nous

croyons que l'effet véritable du plan de sondage dépasse 1 pour l'enquête, en raison des grappes et de la pondération de l'échantillon. (Il est difficile d'estimer l'effet du plan de sondage étant donné le nombre restreint de degrés de liberté.) C'est pourquoi les méthodes de troncation paraissent raisonnables pour cette application, à notre avis.

En raison des degrés de liberté peu nombreux et parce que l'usage de la cocaïne n'est pas rare, on note plus de variations entre les intervalles de confiance à logit, de Breeze et proposés, au tableau 7. Les simulations décrites à la partie 3 nous portent à recommander les intervalles de confiance proposés (avec troncation).

Notre approche pourrait donner l'impression d'être légèrement incohérente dans le cas de cette enquête, en ce sens que nous tolérons des échantillons d'une taille efficace estimée inférieure à la taille de l'échantillon, et les tronquons davantage. Nous pensons toutefois qu'il s'agit d'une approche raisonnablement conservatrice lorsqu'on soupçonne que l'effet véritable du plan de sondage dépasse 1.

5. DISCUSSION

Quoique les intervalles de confiance proposés présentent une couverture probable adéquate pour toutes les simulations ou presque, on ne peut garantir ce résultat pour l'ensemble des configurations de la population (voir le dernier tiers du tableau 5). Il est facile d'illustrer une non-couverture plus importante. Supposons, par exemple, que la population se compose de grappes de 100 éléments et que 10 % des grappes ne renferment que des éléments positifs et les 90 % restants, aucun. En prélevant 10 grappes par échantillonnage aléatoire simple et en prenant toutes les unités des grappes retenues, dans 35 % (= $(1\text{-}0,1)^{10}$) des cas, l'échantillon de 1 000 ne comprendra aucune unité positive. Dans une situation de ce genre, les intervalles proposés se résument aux intervalles binomiaux usuels, c'est à-dire la limite supérieure de l'intervalle de confiance de 95 % est 0,003 (= $1\text{-}0,05^{1/1000}$). Il s'ensuit que la limite supérieure de l'intervalle de confiance de 95 % est inférieure à la valeur véritable de 0,10 dans au moins 35 % des cas, signe d'une sérieuse non-couverture.

Dans les cas d'échantillonnage simples, on peut bâtir des intervalles de confiance qui respectent au moins la probabilité de couverture nominale en tenant compte de tous les agencements possibles de population et en retenant la configuration la moins favorable pour calculer la couverture probable. Dans le cas hypothétique de l'échantillon à un degré mentionné ci-dessus, par exemple, la limite binomiale calculée à partir de 0 unité positive sur 10, soit 0,26 (=$1\text{-}0,05^{1/10}$) pourrait servir de limite supérieure à l'intervalle de confiance de 95 %. Gross et Frankel (1991) ont étudié de tels intervalles, dont le calcul peut devenir très laborieux, et suggèrent eux aussi des approximations réclamant des opérations moins nombreuses.

Comparativement aux approches de ce genre, les intervalles proposés ont pour avantage (1) d'être faciles à calculer, (2) de se plier aux plans d'échantillonnage complexes, y compris à l'ajustement des poids pour la non-réponse et par stratification a posteriori, (3) de généralement garder leur probabilité de couverture nominale, (4) d'être moins conservateurs que les intervalles qui préservent leur couverture nominale sans égard à la configuration de la population et (5) de présenter des propriétés plus intéressantes que les intervalles linéaires, à logit ou de Breeze. Les conclusions (2) et (5) reposent sur les résultats des simulations qui, il va de soi, ne couvrent pas toutes les situations imaginables. Il conviendrait d'entreprendre d'autres recherches en la matière.

REMERCIEMENTS

Les auteurs désirent remercier M. Khare pour leur avoir fourni les estimations sur la prévalence et l'erreur-type attribuable au plan de sondage qui s'y rapporte apparaissant au tableau 6, ainsi que le rédacteur associé et les examinateurs pour leurs précieux commentaires.

BIBLIOGRAPHIE

AGRESTI, A. (1990). *Categorical Data Anlaysis*. New York: Wiley.

BREEZE, E. (1990). *General Household Survey: Report on Sampling Error*. London: Her Majesty's Stationery Office (Office of Population Censuses and Surveys).

CLOPPER, C.J., et PEARSON, E.S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26, 404-413.

COCHRAN, W.G. (1977). *Sampling Techniques*. Troisième édition. New York: Wiley.

FRANKEL, M.R. (1971). *Inference from Survey Samples: An Empirical Investigation*. Ann Arbor, MI: Institute for Social Research, The University of Michigan.

GRAUBARD, B.I., et KORN, E.L. (1996). Modelling the sampling design in the analysis of health surveys. *Statistical Methods in Medical Research*, 5, 263-281.

GROSS, S.T., et FRANKEL, M.R. (1991). Confidence limits for small proportions in complex samples. *Communications in Statistics – Theory and Methods*, 20, 951-975.

GUENTHER,, W.C. (1983). Hypergeometric distributions. Dans *Encyclopedia of Statistical Sciences*, Volume 3, (Éds. S. Kotz et N.L. Johnson). New York: Wiley, 707-712.

JENNINGS, D.E. (1987). How do we judge confidence-interval adequacy? *American Statistician*, 41, 335-337.

JOHNSON, N.L., et KOTZ, S., (1970). *Continuous Univariate Distributions – 2*. New York: Wiley.

JOHNSON, N.L., KOTZ, S. et KEMP, A.W. (1993). *Univariate Discrete Distributions*. Deuxième édition. New York: Wiley.

KORN, E.L., et GRAUBARD, B.I. (1990). Simultaneous testing or regression coefficients with complex survey data: use of Bonferroni *t*-statistics. *American Statistician*, 44, 270-276.

KORN, E.L., et GRAUBARD, B.I. (1998). Variance estimation for superpopulation parameters. *Statistica Sinica*, 8, 1131-1151.

KOVAR, M.G., et JOHNSON, C. (1986). Design effects from the Mexican American portion of the Hispanic Health and Nutrition Examination Survey: a strategy for analysts. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 396-399.

KREWSKI, D., et RAO, J.N.K. (1981). Inference from stratified samples: properties of the linearization, jackknife and balanced repeated replication methods. *Annals of Statistics*, 9, 1010-1019.

McQUILLAN, G.M., KHARE, M., KARON, J.M., SCHABLE, C.A., et VLAHOV, D. (1997). Update on the seroepidemiology of Human Immunodeficiency Virus in the United States household population: NHANES III, 1988-94. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 14, 355-360.

NATIONAL CENTER FOR HEALTH STATISTICS (1985). Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982-84. *Vital and Health Statistics* 1(19). Hyattsville, MD: National Center for Health Statistics.

NATIONAL CENTER FOR HEALTH STATISTICS (1994). Plan and operation of the Third National Health and Nutrition Examination Survey, 1988-94. *Vital and Health Statistics*, 1(32). Hyattsville, MD: National Center for Health Statistics.

PFEFFERMANN, D., et LAVANGE, L. (1989). Regression models for stratified multi-stage cluster samples. Dans *Analysis of Complex Surveys*, (Éds. C.J. Skinner, D. Holt, et T.M.F. Smith). New York: Wiley, 237-260.

SAMHSA (1998). National Household Survey on Drug Abuse: Main Findings 1996. (DHHS Publication No. (SMA) 98-3200). Rockville, MD: SAMHSA.

SAS (1990). *SAS Language: Reference, Version 6.* Première édition. Cary, NC: SAS Institute Inc.

SHAH, B.V., BARNWELL, B.G., et BIELER, G.S., (1995). *SUDAAN User's Manual, Release 6.40.* Research Triangle Park. NC: Research Triangle Institute.

VOLLSET, S.E. (1993). Confidence intervals for a binomial proportion. *Statistics in Medicine*, 12, 809-824