A DIRECT ESTIMATION OF INTRACLUSTER CORRELATION

Jai Won Choi, National Center for Health Statistics, 3700 East West Highway, Hyattsville, MD 20782

1. INTRODUCTION.

Cohen (1960) and Fleiss (1971) developed Kappa statistics of intracluster correlation and measured the agreement between a number of raters when subjects are rated on a nominal scale.

Landis and Koch (1977) introduced the intracluster correlation by random effect model, that is estimated by the components of the analysis of variance. They also measured the agreement or reliability of raters.

These two estimators are asymptotically same when the number of raters for individual subjects are fixed.

However, as mentioned in Fleiss et al (1979), it appears that the the variances of the correlation estimators may give inflated values in both papers. In this note, the direct estimation of intracluster correlation and its closed form of variance are presented, and shows some reduced variances as seen in Table 1.

We will briefly describe the intracluster correlation model in Section 2. Above two estimators of correlation are briefly described, following with a direct estimation and its large sample distribution in Section 3. The variance of direct estimators is given in Appendices A, B, and C. A numerical example is included in Section 4.

Brier (1980) and Kleinman(1973) used method of moments; Cohen (1967) obtained the correlation by the maximum likelihood estimation, while Spearman (1904) and Kendall (1955) used nonparametric approaches, using the ranks of variables.

2. CORRELATION MODEL

Suppose we have a population of A clusters (for example, houses or subjects), $U = (U_1,.,U_i,.., U_A)$, where the cluster i included B_i elementary units (persons in the house or a number of rating on the same subject), expressed as $U_i = (Y_{i1}, ..,Y_{ij},., Y_{iB_i})$. The clusters are indexed by i = 1, ..., A and the units by by j = 1,..., B_i.

Suppose that there are r categories (or ratings) and each elementary unit can be classified into only one of the categories, indexed by h = 1,...,r.

Let $Y'_{ij} = (y_{ij1},...., y_{ijr})$ be the response vector of the (ij)-th person with $\sum_i\sum_j y_{ijh} = 1$ and

$$y_{ijh} = \begin{cases} 1 \text{ if the (ij)-th element falls in cell h} \\ 0\text{, otherwise.} \end{cases} \qquad (1)$$

Let $\sum_i^A \sum_j^{B_i} y_{ijh} = N_h$ be the population counts in the h-th cell and $\sum_h^r N_h = N$ the entire units in the population. Denote the population proportion of the categories by $\pi' = (\pi_1, \pi_2, ..., \pi_r)$, where $\pi_h = N_h/ N$ ($\pi_h > 0$ and $\sum \pi_h = 1$).

Define the pairwise probability for any two members in the same cluster as

$$Pr(y_{ijh} = 1, y_{i'j'h'} = 1) = \begin{cases} \delta_{hh'} & \text{if } i = i' \;\; j \neq j' \text{ and } h \neq h' \\ \delta_{hh} & \text{if } i = i' \;\; j \neq j' \text{ and } h = h' \\ \pi_h & \text{if } i = i' \;\; j = j' \text{ and } h = h' \\ 0 & \text{if } i = i' \;\; j = j' \text{ and } h \neq h' \\ \pi_h \pi_{h'} & \text{if } i \neq i' \end{cases} \qquad (2)$$

where $\delta_{hh'}$ is the probability that one member of the pair from the same cluster falls into the h-th cell and the other into h'-th cell, ($h \neq h'$) and δ_{hh} arises when both elements fall into the h-th cell ($h = h'$).

Note $\sum_{h,h'} \delta_{hh'} = 1$ and, if both members of all pairs fall in the same category, the off-diagonal elements are zeros and the sum of the diagonal elements is one.

Denote the intracluster correlation of any two elements in the same cluster by

$$\rho_{hh'} = \frac{cov(y_{ijh}, y_{ij'h'})}{\sqrt{var(y_{ijh})}\sqrt{var(y_{ij'h'})}} \qquad (3)$$

where $cov(y_{ijh}, y_{ij'h'}) = \delta_{hh'} - \pi_h \pi_{h'}$ and $var(y_{ijh}) = \pi_h(1 - \pi_h)$ from the definition (2).

The ρ_{hh} is the intracluster correlation of the h-th cell when h = h' and the overall intracluster correlation ρ is defined as

$$\rho = \frac{\sum_h (\delta_{hh} - \pi_h^2)}{\sum_h [\pi_h(1 - \pi_h)]} \qquad (4)$$

3. ESTIMATION

3.1. Sample

Denote a sample design [U, S, P], where a one-stage cluster sample S is selected from the population U by the probability P for the cluster sampling. Denote the sample S as

$S = [\ (i,j)\!: i \in S^*, j \in S_i]$

where S* is a sample of "a" clusters, and S_i is a sample of b_i units from the ith cluster. The clusters are indexed by i = 1, ..., a; the units in the ith cluster by j = 1, ..., b_i.

We assume the clusters are independent and the units in the cluster are correlated by a common intracluster correlation ρ_{hh} for the h-th category and ρ for overall categories.

The parameters ρ_{hh} and ρ defined in the models (3) and (4), respectively are estimated from the sample S. The probability P used in the sampling is not involved in our model based estimation.

3.2. Multivariate Analysis of Variance (MANOVA) Estimator

Landis and Koch (1977) used a one-way components of variance model for categorical data to estimate the intracluster correlation in one-stage cluster sample involving an unbalanced design. They estimated the papameters ρ and ρ_{hh} as

$$\hat{\rho} = \frac{\sum_{h=1}^{r} (MS_{ch} - MS_{eh})}{\sum_{h=1}^{r} (MS_{ch} + (d-1)MS_{eh})} \qquad (5)$$

where $d = (n^2 - \sum b_i^2)/n(a-1)$, and $n = \sum b_i$. MS_{ch} and MS_{eh} are the mean squares for the clusters and residual errors with $(a - 1)$ and $(n - a)$ degrees of freedoms, respectively, for the h-th category in a usual MANOVA table. The intracluster correlation $\hat{\rho}_{hh}$ for individual cells is estimated from the (5) by dropping the summation signs as

$$\hat{\rho}_{hh} = \frac{MS_{ch} - MS_{eh}}{MS_{ch} + (d-1)MS_{eh}} \qquad (6)$$

The numerator and denominator of (5) and (6) are unbiased as estimates for those of (3) and (4), respectively.

3.3. Kappa Estimator

Cohen (1960; 1968) used kappa estimator for intracluster correlations. Fleiss (1971) illustrated the use of kappa and weighted kappa with a psychological diagnostic data on 30 patients. Fleiss et al (1979) show the measurement of the extent of agreement beyond chance when for $b_i = b$ (or the number of units in the clusters are the same):

$$\kappa_h = 1 - \frac{\sum_h y_{i+h}(b - y_{i+h})}{ab(b-1)\sum_h \pi_h(1 - \pi_h)}, \qquad (7)$$

where $\sum_j y_{ijh} = y_{i+h}$ and $\sum_h y_{i+h} = b_i$, and the overall measure of agreement is a weighted average of κ as

$$\kappa = \frac{\sum_h \pi_h(1 - \pi_h)\kappa_h}{\sum_h \pi_h(1 - \pi_h)} \qquad (8)$$

The estimators of (7) and (8) are obtained by replacing π_h by $\hat{\pi}_h$, $h = 1,..,r$. It is easy to show $\kappa_h = \rho_{hh}$ if $a/(a-1) \to 1$ for a large a. The variances of (7) and (8) are presented in Fleiss et al (1979).

3.4. Direct Estimator

The estimator of ρ for overall categories can be obtained by direct substitution of $\hat{\pi}_h$ and $\hat{\delta}_{hh}$ in the definition (3) and (4) as

$$\tilde{\rho} = \frac{\sum_h (\hat{\delta}_{hh} - \hat{\pi}_h^2)}{1 - \sum_h \hat{\pi}_h^2} \qquad (9)$$

where $\hat{\pi}_h = \frac{1}{n} \sum_i^a \sum_j^{b_i} y_{ijh}$, $H = \sum_i^a b_i(b_i - 1)$;

and $\hat{\delta}_{hh} = \frac{1}{H} [\sum_i^a y_{i+h}^2 - \sum_i^a \sum_j^{b_i} y_{ijh}^2]$.

"+" means the summation over the corresponding subscript. The $var(\tilde{\rho})$ is given in the Appendix B, (B14).

The estimator of ρ_{hh} for the intracluster correlation of the hth category is obtained by (9) without summation signs as

$$\tilde{\rho}_{hh} = \frac{(\hat{\delta}_{hh} - \hat{\pi}_h^2)}{\hat{\pi}_h(1 - \hat{\pi}_h)}. \qquad (10)$$

The $var(\tilde{\rho}_{hh})$ is given in the Appendix A, (A13). The numerator and denominator of ANOVA estimators $\hat{\rho}$ and $\hat{\rho}_{hh}$ can be expressed as the estimators $\tilde{\rho}$ and $\tilde{\rho}_{hh}$ if $\hat{\delta}_{hh} = \hat{\delta}_{ihh}$ and $\hat{\pi}_h = \hat{\pi}_{ih}$ for all i's. But this is not true in general.

If the nummerator and denominator of $\tilde{\rho}$ and $\tilde{\rho}_{hh}$ are consistent, then the estimators $\tilde{\rho}$ and $\tilde{\rho}_{hh}$ are also consistent since the estimates of a parameter function is consistent as the same function of consistent estimators of the parameters.

However, $\tilde{\rho}$ and $\tilde{\rho}_{hh}$ included biased estimators of numerator and denominator, and we may use the unbiased numerator and denominator, and improve $\tilde{\rho}_{hh}$ and $\tilde{\rho}$ as

$$\hat{\hat{\rho}}_{hh} = \frac{\tilde{\rho}_{hh}(1 - \frac{1}{n}) + \frac{1}{n}}{\frac{\tilde{\rho}_{hh} H}{n^2} + (1 - \frac{H}{n^2})} \qquad (12)$$

with $var(\hat{\hat{\rho}}_{hh}) = var(\tilde{\rho}_{hh})(1 - \frac{1}{n} - \frac{H}{n^2})^2$, (13)

where $var(\tilde{\rho}_{hh})$ is given in (A13) and H in (12).

Note both $\hat{\hat{\rho}}_{hh} \to \tilde{\rho}_{hh}$ and $var(\hat{\hat{\rho}}_{hh}) \to var(\tilde{\rho}_{hh})$ as $n \to \infty$. Following the same steps of the cell intracluster correlations, an unbiased estimator over all categories is obtained as

$$\hat{\hat{\rho}} = \frac{\sum_h^r(\hat{\delta}_{hh} - \hat{\pi}_h^2) + \frac{1}{n}(1 - \sum_h^r \hat{\delta}_{hh})}{(1 - \sum_h^r \hat{\pi}_h^2) - \frac{H}{n^2}(1 - \sum_h^r \hat{\delta}_{hh})} \quad (14)$$

$var(\hat{\hat{\rho}})$ is the same as $var(\tilde{\rho})$ except for the partial derivatives shown in (B2) and (B3), and (C2) and (C3), respectively. With the new partial derivatives (C2) and (C3), we obtain the asymptotic variance $var(\hat{\hat{\rho}})$ as shown in (C4). Note $\hat{\hat{\rho}} \rightarrow \tilde{\rho}$ and $var(\hat{\hat{\rho}}) \rightarrow var(\tilde{\rho})$ for a large n.

Large sample distribution

The theorems on large sample distribution (for example, Cheung, 1969; Bishop, Fienberg, and Holland, 1979) are applied to the above estimators under the usual regularity conditions.

A null hypothesis that specifies the cell correlation ρ_{hh} may be tested by the statistic, for $\rho_{hh} = 0$, (15)

$$Z_h = \frac{\tilde{\rho}_{hh}(1 - 1/n) + 1/n}{[1 - \frac{H}{n^2} + \frac{H}{n^2}\tilde{\rho}_{hh}]\ [1 - \frac{1}{n} - \frac{H}{n^2}]\sqrt{var(\tilde{\rho}_{hh})}}$$

where $Z_h \rightarrow N(0, 1)$ for a large "a". $var(\tilde{\rho}_{hh})$ is given in (A13) and H is a constant multiplier (A12).

A null hypothesis that specifies overall correlation ρ may also be tested by the statistic, for $\rho = 0$,

$$Z = \frac{\sum_h^r(\hat{\delta}_{hh} - \hat{\pi}_h^2) + \frac{1}{n}(1 - \sum_h^r \hat{\delta}_{hh})}{[(1 - \sum_h^r \hat{\pi}_h^2) - \frac{H}{n^2}(1 - \sum_h^r \hat{\delta}_{hh})]\sqrt{var(\hat{\hat{\rho}})}} \quad (16)$$

where $Z \rightarrow N(0, 1)$ for a large "a" or n

4 EXAMPLE

Fleiss (1971) and Landis and koch (1977) used a psychiatric diagnostic data of 30 patients each patient classified separately by six psychiatrists into one of the five response categories :
1) depression, 2) personality disorder,
3) schizophrenia, 4) neurosis, and 5) others.
Here patients are considered as clusters and the six diagnosis as the elements in the cluster. They measured the intracluster correlation and its variances of overall and individual categories in order to test the reliability of the psychiatrists. Table 1 compares the estimates of intracluster correlations and their variances of Landis and Koch (1977) and those of the Fleiss et al (1979) to the direct results.

The three sets of estimates for the intracluster correlations are quite close as expected, and for the direct estimation, the tests of significance for the hypothesis that there is no correlation is rejected at $\alpha = 0.01$ level of significance.

We may test the correlation with a hypothesis of nonzero correlation.

The estimator of MANOVA and direct methods are expected to be the same, and all three results are approximately equal. However, the wide difference of these standard errors requires careful evaluations of these variance estimators.

The variance of MANOVA estimators may be obtained under multinomial assumptions of cell distribution, and compared with the asymptotic variance of the direct estimator.

The further study on the variance of these estimator are required.

REFERENCES

Brier, S. S. (1980). Analysis of Categorical Tables under Cluster Sampling. Biometrika, 67, 591-6.

Cohen, J. A. (1960). Coefficient of agreement for nominal scale. Educational and Psychological Measurement, 20,37-46.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 70, 213-220.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. Psychological Bulletin, 76, 378-82.

Fleiss, J. L., Nee, J. C. M., and J. R. Landis (1979): Large Sample Variance of Kappa in the Case of Different Sets of Raters. Psychological Bulletin, 86, 974-7.

Kendall, M. G. (1955). Rank Corrrlation Methods, 2nd Ed. Hafner, New York.

Kleinmam, J. (1973). Proportion with Extraneous Variances. Single and Independent Samples. Journal of the American Statistical Association, 68, 46-54.

Landis, J. R. and Koch, G. G. (1977). A one-way components of variance model for categorical data. Biometrics, 33, 671-79.

Spearman, C. (1904). The proof and measurement of association between two things. American Journal of Psychology, 15, 72-101.

Appendix A: Variance of $\tilde{\rho}_{hh}$

We only consider $\tilde{\rho}_{hh}$ as the function of $(\hat{\pi}_h, \hat{\delta}_{hh})$ and is not in the from of a direct function of y_{ijh}'s, we will use linear approximation twice to obtain $var(\tilde{\rho}_{hh})$ as shown below.

Denote the partial derivative of $\tilde{\rho}_{hh}$ with respect to $\hat{\pi}_h$ by f_1, and the partial derivative of $\tilde{\rho}_{hh}$ with respective to $\hat{\delta}_{hh}$ by f_2, both evaluated at the parameters π_h and δ_{hh}, assuming that the first order partial derivatives exist.

The linear approximation for $var(\tilde{\rho}_{hh})$ is expressed as

$$var(\tilde{\rho}_{hh}) = [f_1,\ f_2] \begin{vmatrix} var(\hat{\pi}_h) & cov(\hat{\pi}_h, \hat{\delta}_{hh}) \\ cov(\hat{\pi}_h, \hat{\delta}_{hh}) & var(\hat{\delta}_{hh}) \end{vmatrix} \begin{vmatrix} f_1 \\ f_2 \end{vmatrix} \quad (A1)$$

The partial derivatives are given by

$$f_1 = \frac{\partial \tilde{\rho}_{hh}}{\partial \pi_h} = \frac{[(2\pi_h - 1)\delta_{hh} - \pi_h^2]}{(\pi_h(1 - \pi_h))^2} \quad (A2)$$

$$f_2 = \frac{\partial \tilde{\rho}_{hh}}{\partial \delta_{hh}} = \frac{1}{\pi_h(1 - \pi_h)}\ ; \quad (A3)$$

We obtain the $var(\hat{\delta}_{hh})$, $var(\hat{\pi}_h)$, and $cov((\hat{\delta}_{hh}, \hat{\pi}_h)$ by linear approximation for the second time. Since both $\hat{\delta}_{hh}$ and $\hat{\pi}_h$ are now the function of $y_h' = (y_{11h}, y_{12h}, \ldots, y_{nh})$, $(n = \sum_i b_i)$ for the h-th category, and these asymptotic variances and covariance can be obtained by the second linear approximation as

$$\begin{vmatrix} var(\hat{\pi}_h) & cov(\hat{\pi}_h, \hat{\delta}_{hh}) \\ cov(\hat{\pi}_h, \hat{\delta}_{hh}) & var(\hat{\delta}_{hh}) \end{vmatrix} = \begin{vmatrix} J_{\pi h}' \\ J_{\delta h}' \end{vmatrix} W_{hh} [J_{\pi h}\ J_{\delta h}] \quad (A4)$$

where W_{hh} is now the n x n covariance matrix of y_h, obtained from the definition (2), and the n x 1 vector $J_{\pi h}$ is the partial derivatives of $\hat{\pi}_h$ with respect to y_h, and $J_{\delta h}$ is the n x 1 partial derivative vector of $\hat{\delta}_{hh}$ with respect to y_h, both evaluated at the parameters π_h and δ_{hh}. These are

$$J_{\pi h}' = \frac{\partial \hat{\pi}_h'}{\partial y_h} = \frac{1}{n} [\ 1,\ 1,\ \ldots,\ 1],\ \text{respectively.} \quad (A5)$$

$$J_{\delta h}' = \frac{\partial \hat{\delta}_{hh}'}{\partial y_h} = \frac{2\pi_h}{H}[(b_1-1)..(b_1-1),(2_2-1)..(b_2-1),..]. \quad (A6)$$

Substituting (A4) into (A1), we obtain

$$\underset{}{var(\hat{\rho}_{hh})} = \underset{(1x2)}{[f_1\ \ f_2]} \underset{(2xn)}{\begin{vmatrix} J_{\pi h}' \\ J_{\delta h}' \end{vmatrix}} \underset{(nxn)}{W_{hh}} \underset{(nx2)}{[J_{\pi h}\ J_{\delta h}]} \underset{(2x1)}{\begin{vmatrix} f_1 \\ f_2 \end{vmatrix}}, \quad (A7)$$

$$\text{or} = f_1^2 (J_{\pi h} W_{hh} J_{\pi h}) + 2 f_1 f_2 (J_{\pi h} W_{hh} J_{\delta h}) + f_2^2 (J_{\delta h} W_{hh} J_{\delta h}), \quad (A8)$$

where

$(J_{\pi h} W_{hh} J_{\pi h}) = var(\hat{\pi}_h)$, $(J_{\pi h} W_{hh} J_{\delta h}) = cov(\hat{\delta}_{hh}\ \hat{\pi}_h)$,

and $(J_{\delta h} W_{hh} J_{\delta h}) = var(\hat{\delta}_{hh})$; these are:

$$var(\hat{\pi}_h) = \frac{\pi_h(1 - \pi_h)}{n} + \frac{H(\delta_{hh} - \pi_h^2)}{n^2} \quad (A9)$$

$$var(\hat{\delta}_{hh}) = \frac{4\pi_h^2}{H^2} [\pi_h(1 - \pi_h)D + (L - D)(\delta_{hh} - \pi_h^2)] \quad (A10)$$

$$\text{and } cov(\hat{\pi}_h, \hat{\delta}_{hh}) = \frac{2\pi_h}{nH} [\pi_h(1 - \pi_h)H + D(\delta_{hh} - \pi_h^2)] \quad (A11)$$

respectively, and the constant multipliers are

$$H = \sum_i^a b_i(b_i - 1),\ D = \sum_i^a b_i(b_i - 1)^2,\ \text{and } L = \sum_i^a [b_i(b_i - 1)]^2. \quad (A12)$$

Note that the order of convergence to zero of $var(\hat{\pi}_h)$, $var(\hat{\delta}_{hh})$, and $cov(\hat{\pi}_h, \hat{\delta}_{1hh})$ is $o(a^{-1})$. Substituting (A9), (A10), (A11) along with the partial derivatives (A2) and (A3) into (A8), we obtain, using the notation $p_h q_h = \pi_h(1 - \pi_h)$,

$$var(\tilde{\rho}_{hh}) = \frac{[(2\pi_h - 1)\delta_{hh} - \pi_h^2]^2}{[p_h q_h]^4} \left[\frac{p_h q_h}{n} + \frac{H(\delta_{hh} - \pi_h^2)}{n^2}\right]$$
$$+ \frac{2[(2\pi_h - 1)\delta_{hh} - \pi_h^2]}{[p_h q_h]^3} 2\pi_h\left[\frac{p_h q_h}{n} + \frac{D(\delta_{hh} - \pi_h^2)}{nH}\right]$$
$$+ \frac{1}{[p_h p_h]^2} \frac{4\pi_h^2}{H^2} [p_h q_h D + (L - D)(\delta_{hh} - \pi_h^2)] \quad (A13)$$

or using the relationship $\rho_{hh}\ p_h q_h = \delta_{hh} - \pi_h^2$, we may rewrite (A13) as

$$= \frac{[\rho_{hh} + 2\pi_h(1 - \rho_{hh})]^2}{p_h q_h} \frac{[n + \rho_{nn} H]}{n^2}$$
$$- \frac{2\ [\rho_{hh} + 2\pi_h(1 - \rho_{hh})]}{p_h q_h} \frac{2\pi_h[H + \rho_{hh} D]}{nH}$$
$$+ \frac{4\pi_h^2\ [D + (L - D)\rho_{hh}]}{p_h q_h H^2}$$

The order of convergence to zero of $var(\tilde{\rho}_{hh})$ is also $o(a^{-1})$.

Appendix B: Variance of $\tilde{\rho}$ shown in (10)

$\tilde{\rho}$ is now seen as function of $\hat{\theta} = (\hat{\delta}_{11}, \ldots, \hat{\delta}_{rr}$ and $\hat{\pi}_1, \ldots, \hat{\pi}_r)$.
The variance of average intracluster correlation over the r response categories is obtained by linear approximation as

$$var(\tilde{\rho}) = \underset{(1x2r)}{F'} \underset{(2rx2r)}{V} \underset{(2rx1)}{F} \quad (B1)$$

where V is the 2r x 2r covariance matrix of $\hat{\theta}' = (\hat{\delta}_{11}, \ldots, \hat{\delta}_{rr}, \hat{\pi}_1, \ldots, \hat{\pi}_r)$, and $F' = (F_{\delta_1}, \ldots, F_{\delta_r}, F_{\pi_1}, \ldots, F_{\pi_r})$ is 1x2r partial derivative vector of $\tilde{\rho}$ with respect to $\hat{\theta}$, evaluated at $\theta = (\delta_{11}, \ldots \pi_1, \ldots)$; that is,

$$F_{\pi_h} = \frac{\partial\rho}{\partial\pi_h} = \frac{2\pi_h(\sum_h \delta_{hh} - 1)}{[1 - \sum_h^r \pi_h^2]^2}, \quad \text{(B2)}$$

and $$F_{\delta_h} = \frac{\partial\tilde{\rho}}{\partial\delta_{hh}} = \frac{1}{(1 - \sum_h^r \pi_h)}. \quad \text{(B3)}$$

Since the covariance matrix V is now for the 2r estimates, each estimate based on the n variables of $y_{11h}, y_{12h}, \ldots, y_{nh}$ ($n = \sum b_i$) for $h = 1, \ldots, r$, we may use the linear approximation for the second time to find this covariance matrix V = JWJ', where the covariance matrix W is of $y_{11h}, y_{12h}, \ldots, y_{nh}$ for $h = 1, \ldots, r$, expressed as

$$\underset{(nr \times nr)}{W} = \begin{vmatrix} W_{11} & W_{21} & W_{31} & \cdots & W_{r1} \\ W_{12} & W_{22} & W_{32} & \cdots & W_{r2} \\ W_{13} & W_{23} & W_{33} & \cdots & W_{r3} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ W_{1r} & W_{2r} & W_{3r} & \cdots & W_{rr} \end{vmatrix}, \quad \text{(B4)}$$

where each submatrix is the n x n covariance matrix for corresponding category h,h' = 1, ..., r. The submatrices W_{hh} (h = h') on the diagonal are the covariance matrices of $y_{11h}, y_{12h}, \ldots, y_{nh}$ for $h = 1, \ldots, r$, and $W_{hh'}$ ($h \neq h'$) on the off-diagonal are the covariance matrices between $y_{11h}, \ldots, y_{nh}$ and $y_{11h'}, \ldots, y_{nh'}$ for $h \neq h'$ according to the definition (2).

Denote the partial derivatives of $\hat{\pi}$ with respect to y by the (r x nr) diagonal matrix J_π of submatrices J_{π_h} as

$$\underset{(r \times nr)}{J_\pi} = \begin{vmatrix} J_{\pi_1} & 0 & 0 & \cdots & 0 \\ 0 & J_{\pi_2} & 0 & \cdots & 0 \\ 0 & 0 & J_{\pi_3} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & J_{\pi_r} \end{vmatrix} \quad \text{(B5)}$$

where the (1x n) submatrix J_{π_h} represents the partial derivatives of $\hat{\pi}_h$ with respect to y_h, evaluated at π_h (h = 1, ..., r) as shown in (A5), and 0 is the (1 x n) null matrix of zeros.

Denote the partial derivatives of $\hat{\delta}$ with respect to y by

$$\underset{(r \times nr)}{J_\delta} = \begin{vmatrix} J_{\delta_1} & 0 & 0 & \cdots & 0 \\ 0 & J_{\delta_2} & 0 & \cdots & 0 \\ 0 & 0 & J_{\delta_3} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & J_{\delta_r} \end{vmatrix} \quad \text{(B6)}$$

where the (1 x n) submatrix J_{δ_h} represents the partial derivatives of $\hat{\delta}_{hh}$ with respect to y_h, evaluated at π_h (h = 1, ..., r) as shown in (A6).

We now take the linear approximation of the covariance V under usual constraints as

$$V = \begin{vmatrix} J_\pi' \\ J_\delta' \end{vmatrix} W [J_\pi \; J_\delta] = J W J' = \begin{vmatrix} J_\pi' W J_\pi & J_\pi' W J_\delta \\ J_\pi' W J_\delta & J_\delta' W J_\delta \end{vmatrix},$$

(2rx2r)(2rxnr)(nrxnr)(nrx2r) (B7)

where $J' = [J_\pi, J_\delta]$.

Substituting (B7) into (B1), the variance of $\tilde{\rho}$ is given by

$$\text{var}(\tilde{\rho}) = F'(J W J') F \quad \text{(B8)}$$

The double linear estimation is often useful when a direct partial derivative is not easily obtainable. The partial derivation can be repeated further if necessary. The equivalance of the repeated partial derivatives to a single partial derivative can be seen from the fact that

$$\frac{\partial\tilde{\rho}}{\partial y_h} = \frac{\partial\tilde{\rho}}{\partial\hat{\delta}_{hh}} \frac{\partial\hat{\delta}_{hh}}{\partial y_h} \quad \text{(B9)}$$

(B8) can be rewritten explicitly as

$$\text{var}(\tilde{\rho}) = \sum_h^r F_{\pi_h}^2 \text{var}(\hat{\pi}_h) + \sum_{h \neq h'}^r \sum^r F_{\pi_h} F_{\pi_{h'}} \text{cov}(\hat{\pi}_h, \hat{\pi}_{h'})$$

$$+ 2[\sum_h^r F_{\pi_h} F_{\delta_h} \text{cov}(\hat{\delta}_{hh}\; \hat{\pi}_h) + \sum_{h \neq h'}^r \sum^r F_{\pi_h} F_{\delta_{h'}} \text{cov}(\hat{\pi}_h, \hat{\delta}_{h'h'})]$$

$$+ \sum_h^r F_{\delta_h}^2 \text{var}(\hat{\delta}_{hh}) + \sum_{h \neq h'}^r \sum^r F_{\delta_h} F_{\delta_{h'}} \text{cov}(\hat{\delta}_{hh}, \hat{\delta}_{h'h'}) \quad \text{(A10)}$$

where we have seen $\text{var}(\hat{\pi}_h)$, $\text{var}(\hat{\delta}_{hh})$, and $\text{cov}(\hat{\pi}_h, \hat{\delta}_{hh})$ previously in (A9), (A10) and (A11), respectively. We can also obtain the linear approximation of $\text{cov}(\hat{\pi}_h, \hat{\pi}_{h'})$, $\text{cov}(\hat{\delta}_{hh}, \hat{\delta}_{h'h'})$, and $\text{cov}(\hat{\pi}_h, \hat{\delta}_{h'h'})$ for $h \neq h'$ as

$$\text{cov}(\hat{\pi}_h, \hat{\pi}_{h'}) = -\pi_h\pi_{h'}/n + H\,(\delta_{hh'} - \pi_h\pi_{h'})/\,n^2 \quad \text{(B11)}$$

$$\mathrm{cov}(\hat{\pi}_h, \hat{\delta}_{h'h'}) = 2\pi_{h'}[-\pi_h\pi_{h'}H + D(\delta_{hh'} - \pi_h\pi_{h'})]/nH \tag{B12}$$

$$\mathrm{cov}(\hat{\delta}_{1hh}\hat{\delta}_{1h'h'}) = 4\pi_h\pi_{h'}[-D\pi_h\pi_{h'} + (L-D)(\delta_{hh'} - \pi_h\pi_{h'})]/H^2 \tag{B13}$$

where H, D, and L are given in (12)

Substituting (A9), (A10), (A11), (B11), (B12), and (B13), along with partial derivatives (B2) and (B3), $\mathrm{var}(\hat{\rho})$ can be written as,

with the notation $o = 1 - \sum_h^r \pi_h^2$, $\mathrm{var}(\tilde{\rho}) =$

$$\frac{4(\sum_h^r \delta_{hh} - 1)^2}{o^4}\left|\sum_{h=1}^r \pi_h^2 \left[\left[\frac{\pi_h(1-\pi_h)}{n} + \frac{H(\delta_{hh} - \pi_h^2)}{n^2}\right]\right] + \sum_{h\neq h'} \pi_h\pi_{h'}\left[\left[\frac{-\pi_h\pi_{h'}}{n} + \frac{H(\delta_{2hh'} - \pi_h\pi_{h'})}{n^2}\right]\right]\right|$$

$$+ \frac{4(\sum_h^r \delta_{hh} - 1)}{o^3}\left|\sum_h^r \pi_h\left[\left[\frac{2\pi_h[H\pi_h(1-\pi_h) + D(\delta_{hh} - \pi_h^2)]}{nH}\right]\right] + \sum_{h\neq h'}^r \pi_h\left[\left[\frac{2\pi_{h'}[-\pi_h\pi_{h'}H + D(\delta_{hh'} - \pi_h\pi_{h'})]}{nH}\right]\right]\right|$$

$$+ \frac{1}{o^2}\left|\sum_h^r \left[\left[\frac{4\pi_h^2[D\pi_h(1-\pi_h) + (L-D)(\delta_{hh} - \pi_h^2)]}{H^2}\right]\right] + \sum_{h\neq h'}^r\sum^r \left[\left[\frac{4\pi_h\pi_{h'}[-D\pi_h\pi_{h'} + (L-D)(\delta_{hh'} - \pi_h\pi_{h'})]}{H^2}\right]\right]\right|$$

where $D = \sum_i^a b_i(b_i - 1)^2$, $L = \sum_i^a [b_i(b_i - i)]^2$, and $H = \sum_i^a b_i(b_i - 1)$. Note that (B14) approaches zero in the order of $o(a^{-1})$.

Appendix C $\mathrm{var}(\hat{\hat{\rho}})$.

$$\hat{\hat{\rho}} = \frac{U}{D} = \frac{\sum_h^r(\hat{\delta}_{hh} - \hat{\pi}_h^2) + \frac{1}{n}(1 - \sum_h^r \hat{\delta}_{hh})}{(1 - \sum_h^r \hat{\pi}_h^2) - \frac{H}{n^2}(1 - \sum_h^r \hat{\delta}_{hh})} \tag{C1}$$

where U and D are so defined.

The variance of this estimator is the same as the variance of the biased estimator (A14) except the partial derivatives. These new partial derivatives are given by

$$F_{\pi_h} = \frac{\partial\hat{\hat{\rho}}}{\partial\pi_h} = \frac{1}{D^2}\left(1 - \frac{1}{n} - \frac{H}{n^2}\right) 2\pi_h \left(\sum_h^r \delta_{hh} - 1\right) \tag{C2}$$

$$F_{\delta_h} = \frac{\partial\hat{\hat{\rho}}}{\partial\delta_{hh}} = \frac{1}{D^2}\left(1 - \frac{1}{n} - \frac{H}{n^2}\right)\left(1 - \sum_h^r \pi_h^2\right). \tag{C3}$$

Substituting these new partial derivatives into (B10), we obtain

$$\mathrm{var}(\hat{\hat{\rho}}) = \frac{1}{D^4}\left(1 - \frac{1}{n} - \frac{H}{n^2}\right)^2 \times \tag{C4}$$

$$\left\|4(\sum \delta_{hh} - 1)^2\left|\sum_{h=1}^r \pi_h^2\left[\frac{\pi_h(1-\pi_h)}{n} + \frac{H(\delta_{hh} - \pi_h^2)}{n^2}\right] + \sum_{h\neq h'}^r \pi_h\pi_{h'}\left[-\frac{\pi_h\pi_{h'}}{n} + \frac{H(\delta_{hh'} - \pi_h\pi_{h'})}{n^2}\right]\right|\right.$$

$$+4(\sum\delta_{hh} - 1)(1 - \sum \pi_{hh}^2)\left|\sum_{h=1}^r \pi_h\left[2\pi_h\left(\frac{\pi_h(1-\pi_h)}{n} + \frac{D(\delta_{hh} - \pi_h^2)}{nH}\right)\right] + \sum_{h\neq h'}^r \pi_h\left[2\pi_{h'}\left(-\frac{\pi_h\pi_{h'}}{n} + \frac{D(\delta_{hh'} - \pi_h\pi_{h'})}{nH}\right)\right]\right|$$

$$\left. + (1 - \sum \pi_{hh}^2)^2\left|\sum_{h=1}^r\left[4\pi_h^2\left(\frac{\pi_h(1-\pi_h)D}{H^2} + \frac{(L-D)(\delta_{hh} - \pi_h^2)}{H^2}\right)\right] + \sum_{h\neq h'}^r\left[4\pi_h\pi_{h'}\left(\frac{-D\pi_h\pi_{h'}}{H^2} + \frac{(L-D)(\delta_{hh'} - \pi_h\pi_{h'})}{H^2}\right)\right]\right|\right\|$$

Table 1

Intracluster Correlations and their Standard Deviations from the three methods (MANOVA, Direct, kappa)

Category	ESTIMATES				
		Direct			
	MANOVA (SE)	biased (SE)	unbiased (SE)	*Z-value	Kappa (SE)
1	0.254(0.1062)	0.245(0.055)	0.254(0.0532)	*4.780	0.248 (0.1140)
2	0.254(0.0994)	0.245(0.055)	0.254(0.0532)	*4.780	0.248 (0.1140)
3	0.530(0.0719)	0.520(0.132)	0.530(0.1272)	*4.166	0.517 (0.1166)
4	0.481(0.0742)	0.471(0.054)	0.481(0.0525)	*9.165	0.470 (0.1396)
5	0.574(0.1263)	0.566(0.101)	0.576(0.0978)	*5.886	0.565 (0.1277)
Overall	0.440(0.0541)	0.430(0.037)	0.440(0.1072)	*4.108	0.430 (0.0275)

* Significant at $\alpha = 0.01$ level of Z value for testing $\rho = \rho_{hh} = 0$.