VARIANCE OF INTRACLUSTER CORRELATION ESTIMATOR

Jai W. Choi, NCHS, 3700 East-West Hwy, Hyattsville MD 20782

KEY WORDS: Intracluster Correlation, ANOVA Estimator, Variance

1. Introduction

Landis and Koch (1977) used one way random effect model to describe a cluster sample, and presented the ANOVA estimator of intracluster correlation. This is an identical estimator as the Kappa (Cohen, 1977); Fleiss, Nee, and Landis (1979) when the data is balanced. Kraemer (1980), Davies and Fleiss (1982), James (1983), O'Connell and Dobson (1984), and Kempthorne (1982, Unpublished thesis, University of North Carolina) also discussed the intracluster correlation for categorical data.

When the data are distributed as multivariate normal, Anderson (1959) and Searle (1956) present asymptotic variance of intracluster correlation estimator.

The variance of intracluster correlation estimator has been the problem when complex sample survey data such as those collected by the NCHS.

A variance of ANOVA intracluster correlation estimator is presented in this paper.

Following the introduction, Section 2 discusses the variance of intracluster correlation for a single level. Section 3 presents the estimator of overall intracluster correlation with its variance.

1.1 Definition of Intracluster Correlation

Let $Y_{hij} = 1$ if the (ij)-th unit is classified as the h-th category with probability π_h for all i and j and $Y_{hij} = 0$ with probability $1 - \pi_h$.

We use the subscripts h (h = 1, ..., r) for r response categories for the sample of "a" clusters indexed by i (i = 1, ..., a) and the ith cluster includes b_i units indexed by j (j = 1,..., b_i).

We assume that the clusters are independent. But the units in the cluster are correlated by ρ_h in the level h and by ρ for overall categories and that a probability sample of two stages is taken with replacement.

Let Y_{hij} be a random variable, discrete, which is expressed as

$$y_{hij} = u_h + c_{hi} + e_{hij} \qquad (1)$$

where c_{hi} and e_{hij} are with mean zeros and variance σ^2_{hc} and σ^2_{he}, respectively. The variables c_{hi} and e_{hij} are the random samples of size "a" and "n" ($n = \Sigma b_i$) from these two populations with not assumptions on the distribution. But c_{hi}'s and e_{hij}'s are assumed uncorrelated.

We use the notations c_h and e_h for σ^2_{hc} and σ^2_{he}.

From this model, $E(y_{hij}) = u_h$ and $V(y_{hij}) = c_h + e_h$ for all i and j where E and V are the expectation and variance operators.

Another expression for discrete data could be $c_h = \rho_h \pi_h(1 - \pi_h)$ and $e_h = (1 - \rho_h) \pi_h(1 - \pi_h)$, where ρ_h is a positive common intracluster correlation for the cluster for the h-th category.

The intracluster correlation for the h-th category is defined as

$$\rho_h = \frac{c_h}{c_h + e_h} \qquad (2)$$

and the intracluster correlation over all categories is defined as

$$\rho = \frac{c_+}{c_+ + e_+} \qquad (3)$$

where $c_+ = \Sigma c_h$ and $e_+ = \Sigma e_h$.

1.2 Estimator of ρ_h

The estimators of ρ_h and ρ are previously presented in terms of the ANOVA sums of squares for within and between clusters (Landis and Koch, 1977).

Searle (1956) wrote it in terms of the least square estimators of e_h and c_h. We can express the estimator of (2) as equation (4) below, this estimator is the same as that of Landis et al and Searle estimator, but in different form.

We can rewrite the ANOVA estimator

$$\hat{\rho}_h = \frac{\hat{U}_h}{\hat{D}_h} = \frac{\sum_i^a [c_{1i} T_{1hi} + c_{2i} T_{2hi}]}{\sum_i^a [c_{3i} T_{1hi} + c_{4i} T_{2hi}]} \qquad (4)$$

where $c_i = (n - 1)/(a-1)b_i$,

$$c_{1i} = (c_i - 1)/(d(n-a))$$

$$d = (n^2 - \Sigma b_i^2)/[n(a - 1)].$$

$$c_{2i} = c_i/(d(n-a))$$

$$c_{3i} = \{c_i - d(1/b_i - 1) - 1)\}/(d(n-a))$$

$$c_{4i} = (c_i - d/b_i)/(d(n-a))$$

$$T_{1hi} = \Sigma_{j=1}^{b_i} a^2_{hij},$$

$$T_{2hi} = \Sigma_{j \neq j'}^{b_i} a_{hij} \; a_{hij'},$$

where $a_{hij} = (y_{hij} - \bar{y}_h)$,

We assume that $(\bar{y}_h - \bar{Y}_h) \to 0$ so that we can replace the sample mean $\bar{y}_h$ with the population mean $\bar{Y}_h$ in the derivation of its variance below.

2. Variance of $\hat{\rho}_h$

The asymptotic variance of (4) can be obtained by delta method as

$$V(\hat{\rho}_h) = G_h' \; V_h \; G_h \qquad (5)$$

where $G_h' = (1/D_h, \; -U_h/D_h^2)$, partial derivative vector, of $\hat{\rho}_h$ with respect to $\hat{U}_h$ and $\hat{D}_h$, evaluated at the (U_h, D_h). The variance covariance matrix of $\hat{U}_h$ and $\hat{D}_h$ is V_h with variances $V(\hat{U}_h)$ and $V(\hat{D}_h)$ on the diagonal and the covariance $C(\hat{U}_h, \hat{D}_h)$ on the off-diagonal. Thus, we can rewrite (5) as

$$V(\hat{\rho}_h) = [1/D_h, -U_h/D_h^2] \begin{vmatrix} V(\hat{U}_h) & C(\hat{U}_h \hat{D}_h) \\ C(\hat{U}_h, \hat{D}_h) & V(\hat{D}_h) \end{vmatrix} \begin{vmatrix} 1/D_h \\ -U_h/D_h^2 \end{vmatrix} \qquad (6)$$

or

$$V(\hat{\rho}_h) = \sum_i^a \frac{1}{D_h^2} [(c_{1i} - R_h c_{3i})^2 \; V(T_{1hi}) \qquad (7)$$
$$+ (c_{2i} - R_h c_{4i})^2 \; V(T_{2hi})$$
$$+ (c_{1i} - R_h c_{3i})(c_{2i} - R_h c_{4i}) \; C(T_{1hi}, T_{2hi})]$$

where $R_h = U_h/D_h$. The variances $V(T_{1hi})$, $V(T_{2hi})$, and covariance $C(T_{1hi}, T_{2hi})$ are shown in Appendix 1.

3. Estimator of ρ

The overall intracluster correlation estimator can be written as

$$\hat{\rho} = \frac{\sum_h^r \hat{U}_h}{\sum_h^r \hat{D}_h} = \frac{\hat{U}_+}{\hat{D}_+} \text{ (say)} \qquad (8)$$

where $\hat{U}_h$ and $\hat{D}_h$ are already defined in Section 2. The sign "+" means the sum over the subscript h.

3.1 Variance of $\hat{\rho}$

The variance of intracluster correlation over all cells can be obtained by the same method as by that of a single cell. The first order approximation of the variance is given as

$$V(\hat{\rho}) = G'V\,G \qquad (9)$$

with $G' = (G_1, G_2, \ldots, G_r)$, where $G_h = (1/D, -U/D^2)$ for h=1,...,r and the covariance matrix V includes the submatrices V_{hh} on the diagonal and submatrices $V_{hh'}$ on the off-diagonal.

The submatrix V_{hh} has the variances $V(\hat{U}_h)$ and $V(\hat{D}_h)$ on the diagonal and covariance $C(\hat{U}_h, \hat{D}_h)$ on the off-diagonal, while the submatrix $V_{hh'}$ includes the covariance $C(\hat{U}_h, \hat{U}_{h'})$ and $C(\hat{D}_h, \hat{D}_{h'})$ on the diagonal $C(\hat{U}_h, \hat{D}_{h'})$ on the off-diagonal.

The equation (9) is rewritten as

$$V(\hat{\rho}) = (G_1, \ldots, G_r) \begin{vmatrix} V_{11} & V_{12} & \cdots & V_{1r} \\ V_{21} & V_{22} & \cdots & V_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ V_{r1} & V_{r2} & \cdots & V_{rr} \end{vmatrix} \begin{vmatrix} G_1 \\ G_2 \\ \vdots \\ G_r \end{vmatrix} \qquad (10)$$

$$V(\hat{\rho}) = \sum_{h=1}^r G_h' \, V_{hh} \, G_h + \sum_{h \neq h'}^r G_h' \, V_{hh'} \, G_{h'} \qquad (11)$$

where G_h is the partial derivative vector of $\hat{\rho}$ with respect to $\hat{U}$ and $\hat{D}$, evaluated at the $U = (U_1, \ldots, U_r)$ and $D = (D_1, \ldots, D_r)$ under the usual assumptions of first order approximation of ratio estimate. We can rewrite (11) as

$$V(\hat{\rho}) = \frac{1}{D_+^2} [\sum_h^r \{V(\hat{U}_h) - 2\,R \;\; C(\hat{U}_h \hat{D}_h) + R^2 \; V(\hat{D}_h)\} \qquad (12)$$
$$+ \sum_{h \neq h'}^r \{C(\hat{U}_h \hat{U}_{h'}) - R\,C(\hat{U}_{h'}, \hat{D}_h)$$
$$- R\,C(\hat{U}_h \hat{D}_{h'}) + R^2 \; C(\hat{D}_h \hat{D}_{h'})\}]$$

where $V(\hat{U}_h)$, $V(\hat{D}_h)$, and $C(\hat{U}_h, \hat{D}_h)$ are shown before in Appendix 1.

The covariances $C(\hat{U}_h, \hat{U}_{h'})$, $C(\hat{D}_h, \hat{D}_{h'})$, and $C(\hat{U}_h, \hat{D}_{h'})$, and the final form of (12) is shown in the Appendix 2.

4. Comments

Applications to actual data are needed to see if these formulas are reasonable.

5. References

T. W. Anderson (1959) An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc.

M. Davies and J. L. Fleiss (1982). Measurement Agreement for Multinomial Data. Biometrics 38, 1047-51. 76-77.

Joseph L. Fleiss, John C. M. Nee, J. Richard Landis (1979). Large Sample Variance of Kappa in the Case of Different Sets of Raters. Psychological Bulletin, Vol. 86, No. 5. 974-977.

I. R. James (1983). Analysis of Nonagreements among Multiple Raters. Biometrics 39, 651-57

J. Richard Landis and Gary G. Koch (1977). A one way components of variance model for categorical data. Biometrics 33, 671-79.
H. C Kraemer (1980) Extension of the Kappa Coefficient. Biometrics 36, 207-216.
D. L. O'Connell and Annette J. Dobson (1984). General Observer-Agreement Measures in Individual Subjects and Groups of Subjects. Biometrics 40, 973-83.
P. D. Oldham (1980) A Note on the Analysis of Repeated Measurements of the Same Subjects. Journal of Chronic Disease Vol. 15, 969-77.
S. R. Searle (1956) Matrix methods in components of variance and covariance analysis. Annals of Mathematical Statistics. Vol. 27, 737-748.

Appendix 1

We derive the variances $V(\hat{U}_h)$ and $V(\hat{D}_h)$, and covariance $C(\hat{U}_h, \hat{D}_h)$ as

$$V(U_h) = \sum_i^a [c_{1i}^2 V(T_{1hi}) + 2 c_{1i} c_{2i} C(T_{1hi}, T_{2hi}) + c_{2i}^2 V(T_{2hi})], \quad (a1)$$

$$V(D_h) = \sum_i^a [c_{3i}^2 V(T_{1hi}) + 2 c_{3i} c_{4i} C(T_{1hi}, T_{2hi}) + c_{4i}^2 V(T_{2hi})], \quad (a2)$$

$$C(U_h, D_h) = \sum_i^a [c_{1i} c_{3i} V(T_{1hi}) + (c_{1i} c_{4i} + c_{2i} c_{3i}) C(T_{1hi}, T_{2hi}) + c_{2i} c_{4i} V(T_{2hi})]. \quad (a3)$$

where the variances $V(T_{1hi})$, $V(T_{2hi})$, and covariance and $C(T_{1hi}, T_{2hi})$ are shown below as

$$V(T_{1hi}) = b_i \sigma_h^4 + b_i(b_i - 1) \sigma_h^{22} - b_i^2 (\sigma_h^2)^2 \quad (a4)$$

$$V(T_{2hi}) = 2b_i(b_i - 1)\sigma_h^{22} + 4b_i(b_i - 1)(b_i - 2)\sigma_h^{211} + b_i(b_i-1)(b_i-2)(b_i-3)\sigma_h^{1111} - \{b_i(b_i - 1)\sigma_h^{11}\}^2, \quad (a5)$$

$$C(T_{1hi}, T_{2hi}) = 2 b_i(b_i - 1) \sigma_h^{13} + b_i(b_i-1)(b_i-2)\sigma_h^{211} - b_i^2(b_i-1)\sigma_h^2 \sigma_h^{11}. \quad (a6)$$

The cross product moments σ_h^2, σ_h^4, σ_h^{11}, σ_h^{22}, σ_h^{31}, σ_h^{211}, and σ_h^{1111} are defined as following.

$$E(a_{hij}^s) = \sum_i^A \sum_j^{B_i} \frac{a_{hij}^s}{A B_i} = \sigma_h^s \quad (s=2 \text{ or } 4) \quad (a7)$$

$$E(a_{hij}^s a_{hij'}^t) = \sum_i^A \sum_{j \neq j'}^{B_i} \frac{a_{hij}^s a_{hij'}^t}{AB_i(B_i-1)} = \sigma_h^{st} \quad (a8)$$

where (s,t) = (1,1) or (2,2) or (3,1),

$$E(a_{hij}^2 a_{hij'}^1 a_{hij''}^1) = \sum_i^A \sum_{j \neq j \neq j'}^{B_i} \frac{a_{hij}^2 a_{hij'} a_{hij''}}{AB_i(B_i-1)(B_i-2)} = \sigma_h^{211} \quad (a9)$$

$$E(a_{hij} a_{hij'} a_{hij''} a_{hij'''}) = \sum_i^A \sum_{j \neq j' \neq \ldots j'''}^{B_i} \frac{a_{hij} a_{hij'} a_{hij''} a_{hij'''}}{AB_i(B_i-1)(B_i-2)(B_i-3)} = \sigma_h^{1111} \text{ (say)}. \quad (a10)$$

We should have at least four members in a cluster for the existence of the fourth cross product moment. For a cluster of less than four members, a cross product of four or more members does not exist.

Different results can be obtained, depending on how we define the cross product moments in the above equations. For instance, these may be defined by a probability model.

A set of unbiased estimates of above cross product moments are

$$\hat{\sigma}_h^s = \sum_i^a \sum_j^{b_i} \frac{a_{hij}}{ab_i} \quad (a11)$$

$$\hat{\sigma}_h^{st} = \sum_i^a \sum_{j \neq j'}^{b_i} \frac{a_{hij}^s a_{hij'}^t}{ab_i(b_i-1)} \quad (a12)$$

$$\hat{\sigma}_h^{211} = \sum_i^a \sum_{j \neq j' \neq j''}^{b_i} \frac{a_{hij}^2 a_{hij'} a_{hij''}}{ab_i(b_i-1)(b_i-2)} \quad (a13)$$

$$\hat{\sigma}_h^{1111} = \sum_i^a \sum_{j \neq j' \neq j'' \neq j'''}^{b_i} \frac{a_{hij} a_{hij'} a_{hij''} a_{hij'''}}{ab_i(b_i-1)(b_i-2)(b_i-3)} \quad (a14)$$

We can rewrite above expressions, using the notation $\sum a_{hij}^c = a_{hi+}^c$ for any positive integer c.

$$\sum_{j \neq j'}^{b_i} a_{hij} a_{hij'} = (a_{hi+})^2 - a_{hi+}^2 \quad (a15)$$

$$\sum_{j \neq j'}^{b_i} a_{hij}^2 a_{hij'} = a_{hi+}^2 a_{hi+} - a_{hi+}^3 \quad (a16)$$

$$\sum_{j\neq j'}^{b_i} a^3_{hij}\, a_{hij'} = a^3_{hi+}\, a_{hi+} - a^4_{hi+} \quad (a17)$$

$$\sum_{j\neq j'}^{b_i} a^2_{hij}\, a^2_{hij'} = (a^2_{hi+})^2 - a^4_{hi+} \quad (a18)$$

$$\sum_{j\neq j'\neq j''}^{b_i} a_{hij}\, a_{hij'}\, a_{hij''} \quad (a19)$$

$$= (a_{hi+})^3 + 2a^3_{hi+} - 3a^2_{hi+}\, a_{hi+}$$

$$\sum_{j\neq j'\neq j''}^{b_i} a^2_{hij}\, a_{hij'}\, a_{hij''} \quad (a20)$$

$$= a^2_{hi+}(a_{hi+})^2 - 2a_{hi+}\, a^3_{hi+} + 2\, a^4_{hi+} - (a^2_{hi+})^2$$

$$\sum_{j\neq j'\neq j''\neq j'''}^{b_i} a_{hij}\, a_{hij'}\, a_{hij''}\, a_{hij'''} \quad (a21)$$

$$= (\sum_j^{b_i} a_{hij})^4 - \sum_j^{b_i} a^4_{hij}$$

$$- 3 \sum_{j\neq j'}^{b_i} a^2_{hij}\, a^2_{hij'} - 4 \sum_{j\neq j}^{b_i} a^3_{hij}\, a_{hij'}$$

$$- 6 \sum_{j\neq j'\neq j''}^{b_i} a^2_{hij}\, a_{hij'}\, a_{hij''}.$$

$$= (a_{hi+})^4 - 6a^4_{hi+} + 3(a^2_{hi+})^2 + 8\, a_{hi+}\, a^3_{hi+} - 6a^2_{hi+}(a_{hi+})^2$$

From (a17)-(a21), the computation of cross product moments are more manageable than the original form.

Appendix 2

For $h \neq h'$, we can write the covariances

$$C(\hat{U}_h, \hat{U}_{h'}) = \sum_i^a [c^2_{1i}\, C(T_{1hi} T_{1h'i}) \quad (b1)$$

$$+ c_{1i}\, c_{2i}\, \{C(T_{1h'i}\, T_{2hi}) + C(T_{1hi}\, T_{2h'i})\}$$

$$+ c^2_{2i}\, C(T_{2hi}\, T_{2h'i})]$$

$$C(\hat{D}_h, \hat{D}_{h'}) = \sum_i^a [c^2_{3i} C(T_{1hi}\, T_{1h'i}) \quad (b2)$$

$$+ c_{3i}\, c_{4i}\{C(T_{1h'i}\, T_{2hi}) + C(T_{1hi}\, T_{2h'i})\}$$

$$+ c^2_{4i}\, C(T_{2hi}\, T_{2h'i})]$$

$$C(\hat{U}_h, \hat{D}_{h'}) = \sum_i^a [c_{1i} c_{3i} C(T_{1hi}\, T_{1h'i}) \quad (b3)$$

$$+ c_{2i} c_{4i}\, C(T_{2hi}, T_{2h'i})$$

$$+ c_{1i}\, c_{4i}\, C(T_{1hi} T_{2h'i}) + c_{2i} c_{3i} C(T_{1h'i} T_{2hi})].$$

Using the previous results of $V(U_h)$, $V(D_h)$, and $C(U_h, D_h)$ in Appendix 1, and above (b1), (b2), and (b3), we can rewrite the variance (12) as

$$V(\hat{\rho}) = \frac{1}{D^2_+} \sum_i^a [\, (c_{1i} - Rc_{3i})^2 \{\sum_h^r V(T_{1hi}) \quad (b4)$$

$$+ \sum_{h\neq h'}^r C(T_{1hi}\, T_{1h'i})\}$$

$$+ (c_{1i} - Rc_{3i})(c_{2i} - Rc_{4i})\{2 \sum_h^r C(T_{1hi}\, T_{2hi})$$

$$+ \sum_{h\neq h'}^r (C(T_{1hi} T_{2h'i}) + C(T_{1h'i} T_{2hi})\,)\}$$

$$+ (c_{2i} - Rc_{4i})^2 \{\sum_h^r V(T_{2hi}) + \sum_{h\neq h'}^r C(T_{2hi}\, T_{2h'i})\}\,]$$

where the form of covariance between T_{1hi} and $T_{1h'1'}$, T_{2hi} and $T_{2h'i'}$, or T_{1hi} and $T_{2h'i}$ are obtained as

$$C(T_{1hi}, T_{1h'i}) = b_i(b_i - 1)\, E(a^2_{hij}\, a^2_{h'ij'}) \quad (b5)$$

$$- b^2_i\, E(a^2_{hij})\, E(a^2_{h'ij})$$

$$C(T_{1hi}, T_{2h'i}) = \quad (b6)$$

$$= b_i(b_i - 1)\,(b_i - 2)\, E(a^2_{hij}\, a_{h'ij'}\, a_{h'ij''})$$

$$- b^2_i\, (b_i - 1)\, E(a^2_{hij})\, E(a_{h'ij}\, a_{h'ij'}),$$

$$C(T_{2hi}, T_{2h'i}) = \quad (b7)$$

$$= b_i(b_i - 1)(b_i - 2)(b_i - 3)\, E(a_{hij}\, a_{hij'}\, a_{h'ij''} a_{h'ij'''})$$

$$- b^2_i\, (b_i - 1)^2\, E(a_{hij}\, a_{hij'})\, E(a_{h'ij}\, a_{h'ij'}).$$

$C(T_{1h'i}, T_{2hi})$ is the same as (b6) except h and h' exchanged.

where the expected values of cross products are defined as

$$E(a^s_{hij}\, a^t_{h'ij'}) = \sum_i^A \sum_{j\neq j'}^{B_i} \frac{a^s_{hij}\, a^t_{h'ij'}}{AB_i(B_i - 1)} = \sigma^{st}_{hh'} \quad (b8)$$

(s,t) = (2,2) or (1,1) in (b5),

$$E(a^2_{hij}\, a_{h'ij'}\, a_{h'ij''}) \quad (b9)$$

$$= \sum_i^A \sum_{j\neq j'\neq j''}^{B_i} \frac{a^2_{hij}\, a_{h'ij'}\, a_{h'ij''}}{AB_i(B_i - 1)(B_i - 2)} = \sigma^{211}_{hh'h'}$$

as seen in (b6),

$$E(a_{hij}\, a_{hij'}\, a_{h'ij''}\, a_{h'ij'''}) \qquad \text{(b10)}$$

$$= \sum_{i}^{A} \sum_{j\neq j'\neq \dots j'''}^{B_i} \frac{a_{hij}\; a_{hij'}\; a_{h'ij''}\; a_{h'ij'''}}{A\; B_i\; (B_i-1)(B_i-2)(B_i-3)}$$

$$= \sigma^{1111}_{hhh'h'} \text{ (say)}$$

as seen in (b7).

These expected values may be estimated from the sample. The units in a cluster may be distributed into categories. Let the number of unit (or units) of the ith cluster falling into the hth cell be b_{hi} so that the sum of b_{hi} over all cells is b_i. Since one unit can belong to only one cell, the estimates can be written as

$$\hat{\sigma}^{st}_{hh'} = \sum_{i}^{a} \sum_{j}^{b_{hi}} \sum_{j'}^{b_{h'i}} \frac{a^{s}_{hij}\; a^{t}_{h'ij'}}{a\; b_{hi}\; b_{h'i}} \qquad \text{(b11)}$$

$$\hat{\sigma}^{211}_{hh'h'} = \sum_{i}^{a} \sum_{j=1}^{b_{hi}} \sum_{j'\neq j''}^{b_{h'i}} \frac{a^{2}_{hij}\; a_{h'ij'}\; a_{h'ij''}}{a\; b_{hi}\; b_{h'i}(1-b_{h'i})} \qquad \text{(b12)}$$

$$\hat{\sigma}^{1111}_{hhh'h'} = \qquad \text{(b13)}$$

$$= \sum_{i}^{a} \sum_{j\neq j'}^{b_{hi}} \sum_{j''\neq j'''}^{b_{h'i}} \frac{a_{hij}\; a_{hij'}\; a_{h'ij''}\; a_{h'ij'''}}{ab_{hi}(b_{hi}-1)\; b_{h'i}(1-b_{h'i})}$$

Note that the following results may be used to rewrite (b11), (b12), and (b13) for easier computation.

$$\sum_{j}^{b_i} \sum_{j'\neq j''}^{b_i} a^{2}_{hij}\; a_{h'ij'} a_{h'ij''}$$

$$= (\sum_{j}^{b_{hi}} a^{2}_{hij})\{(\sum_{j\neq j'}^{b_{h'i}} a_{h'ij})^{2} - \sum_{j}^{b_{h'i}} a^{2}_{h'ij}\}$$

$$= \dot{a}^{2}_{hi+}((\dot{a}_{h'i+})^{2} - \dot{a}^{2}_{h'i+})$$

where $\dot{a}_{hi+}$ is the sum of the subscript j over the b_{hi} elements.

$$\sum_{j\neq j}^{b_i} \sum_{j''\neq j'''}^{b_i} a_{hij}\; a_{hij'}\; a_{h'ij''}\; a_{h'ij'''}$$

$$= (\sum_{j\neq j'}^{b_{hi}} a_{hij} a_{hij'})(\sum_{j\neq j'}^{b_{h'i}} a_{h'ij}\; a_{h'ij'})$$

$$= [(\dot{a}_{hi+})^{2} - \dot{a}^{2}_{hi+}][(\dot{a}^{2}_{h'ij}) - \dot{a}^{2}_{h'i+}]$$

Using these results, we can write (b11) - (b13) as

$$\hat{\sigma}^{st}_{hh'} = \sum_{i}^{a} \frac{a^{s}_{hi+}\; a^{t}_{h'i+}}{b_{hi}\; b_{h'i}} \qquad \text{(b14)}$$

$$\hat{\sigma}^{211}_{hh'h'} = \sum_{i}^{a} \frac{\dot{a}^{2}_{hi+}\{(\dot{a}_{h'i+})^{2} - \dot{a}^{2}_{h'i+}\}}{a\; b_{hi}\; b_{h'i}(1-b_{h'i})} \qquad \text{(b15)}$$

$$\hat{\sigma}^{1111}_{hhh'h'} = \qquad \text{(b16)}$$

$$= \sum_{i}^{a} \frac{\{(\dot{a}_{ih+})^{2} - \dot{a}^{2}_{ih+}\}\{(\dot{a}_{ih'+})^{2} - \dot{a}^{2}_{ih'+}\}}{a\; b_{hi}\; (1 - b_{hi})\; b_{h'i}(1 - b_{h'i})}$$

Using sample estimators, we can now obtain the variance estimator from the variance formula (b4).

Note that $E(\dot{a}^{c}_{hij}) = \sum_{j}^{B_{hi}} a_{hij}/ B_{hi}$.