Ann. Inst. Statist. Math.
Vol. 42, No. 3, 561–580 (1990)

INFERENCES ON INTERCLASS AND INTRACLASS CORRELATIONS IN MULTIVARIATE FAMILIAL DATA*

SADANORI KONISHI[1, 2,]** AND C. G. KHATRI[2,]***

[1] The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan
[2] Center for Multivariate Analysis, Department of Statistics, The Pennsylvania State University, University Park, PA 16802, U.S.A.

(Received May 18, 1989; revised November 21, 1989)

Abstract. Inference procedures for interclass and intraclass correlations are given in the multivariate context of familial data for which measurements are taken on more than one characteristic. Unified estimators are proposed based on a certain class of unbiased estimators of covariance matrices. Asymptotic distributions of the proposed estimators are derived under the assumption of multivariate normality. The results can be used to construct approximate confidence intervals and test procedures.

Key words and phrases: Asymptotic distribution, interclass and intraclass correlation matrices, interval estimation, hypothesis testing, multivariate familial data.

1. Introduction

In the biological research on genetics, the interclass and intraclass correlations play an important role in estimating the degree of resemblance among family members with respect to some characteristics, such as blood pressure, cholesterol or lung capacity. Inference procedures have been constructed based on a sample of families, each of which consists of the parent's score and several siblings' scores on a single characteristic (see, for example, Donner (1986) and Srivastava and Keen (1988)).

Very little work has been done on statistical inferences concerning interclass and intraclass correlations in the multivariate context of familial data, for which measurements are taken on several characteristics. This occurs in the analysis of multivariate familial data where one may be interested in assessing the interrelationships among different characteristics.

---

*Research supported by the Department of Statistics, the Pennsylvania State University and the Air Force Office of Scientific Research under Grant AFSO-88-0030.

**Now at the Institute of Statistical Mathematics.

***On leave from Gujarat University, Ahmedabad 380009, India.

The multivariate generalization of intraclass correlation has been done by Rao (1945, 1953), when families have the same number of siblings. In the case where the number of siblings varies among families, Srivastava *et al.* (1988) first considered the problem of estimating the familial correlations in the multivariate situation, and obtained the asymptotic variances of the proposed estimators.

In this paper, inference procedures for interclass and intraclass correlations are discussed in the multivariate situation of more than one characteristic. Unified estimators are proposed based on a certain class of unbiased estimators of covariance matrices. The resultant estimators include the multivariate extension of previous estimators—the pairwise, ensemble (Rosner *et al.* (1977)) and Srivastava's (1984) estimators of the interclass correlation, the analysis of variance (Fisher (1958)), pairwise and a weighted pairwise (Karlin *et al.* (1981)) estimators of the intraclass correlation, etc. Unified formulae to asymptotic distributions of the proposed estimators are derived under the assumption of multivariate normality. The asymptotic results can be used to construct approximate confidence intervals and test procedures for interclass and intraclass correlations.

In the univariate situation of a single characteristic, various kinds of estimators of the interclass and intraclass correlations have been proposed as alternatives to the maximum likelihood estimator which can be evaluated only numerically. Several comparisons have been made of the large-sample and finite-sample properties of these estimators. A survey of work for inferences concerning the intraclass correlation was given by Donner (1986). It has been shown through theoretical and simulation studies that inferences based on the pairwise, ensemble and Srivastava's estimators are comparable or compare favourably with those based on the maximum likelihood estimator (Rosner (1979), Konishi (1982, 1985), Donner and Bull (1984), Srivastava (1984), Srivastava and Katapa (1986), Donner and Eliasziw (1988), Srivastava *et al.* (1988), Srivastava and Keen (1988)). It is therefore of interest to extend these estimators to the multivariate situation of several characteristics.

2. The model for multivariate familial data

Suppose we have a random sample of N families, each of which has a different number of offspring. Let

$$z_{(\alpha)} = (y'_\alpha, x'_{1\alpha}, x'_{2\alpha}, \ldots, x'_{k_\alpha,\alpha})'; \qquad \alpha = 1, 2, \ldots, N$$

denote measurements on the α-th family, where $y_\alpha = (y_{1\alpha}, y_{2\alpha}, \ldots, y_{p\alpha})'$ is the parent's score on p characteristics, and $x_{j\alpha} = (x_{1j,\alpha}, x_{2j,\alpha}, \ldots, x_{qj,\alpha})'$ the score of the j-th child of the k_α (≥ 1) siblings on q characteristics. Let

$$E[y_\alpha] = \mu_m, \qquad E[x_{j\alpha}] = \mu_s \quad \text{for} \quad j = 1,\ldots,k_\alpha$$

and

$$\text{(2.1)} \qquad \begin{aligned} &\operatorname{cov}(y_\alpha) = \Sigma_m = (\sigma_{ab}^{(m)}), && \operatorname{cov}(x_{j\alpha}) = \Sigma_s = (\sigma_{ab}^{(s)}), \\ &\operatorname{cov}(y_\alpha, x_{j\alpha}) = \Sigma_{ms} = (\sigma_{ab}^{(ms)}), && \operatorname{cov}(x_{i\alpha}, x_{j\alpha}) = \Sigma_{ss} = (\sigma_{ab}^{(ss)}) \end{aligned}$$

for $i \neq j = 1, 2,\ldots, k_\alpha$.

This implies that each family may have different numbers of siblings, and that there is no difference among siblings with regard to the characteristics under consideration. It is assumed that $z_{(\alpha)}$ $(\alpha = 1,\ldots, N)$ follows a $(p + qk_\alpha)$-variate normal distribution with mean vector $\mu_{(\alpha)} = (\mu_m', \mu_s',\ldots, \mu_s')'$ and covariance matrix $\Sigma_{(\alpha)}$ having the structure

$$\text{(2.2)} \qquad \Sigma_{(\alpha)} = \begin{bmatrix} \Sigma_m & e_{k_\alpha}' \otimes \Sigma_{ms} \\ e_{k_\alpha} \otimes \Sigma_{ms}' & I_{k_\alpha} \otimes \Sigma_s + (e_{k_\alpha} e_{k_\alpha}' - I_{k_\alpha}) \otimes \Sigma_{ss} \end{bmatrix}$$

where $e_{k_\alpha} = (1, 1,\ldots, 1)'$ is the k_α-dimensional vector, I_{k_α} is the identity matrix of order k_α and $A \otimes B$ denotes the Kronecker product of the matrices A and B.

For the diagonal elements of Σ_m and Σ_s in (2.1), let

$$D_{(\alpha)} = \operatorname{diag}\,[\sigma_{11}^{(m)},\ldots, \sigma_{pp}^{(m)}, \sigma_{11}^{(s)},\ldots, \sigma_{qq}^{(s)},\ldots, \sigma_{11}^{(s)},\ldots, \sigma_{qq}^{(s)}],$$

a $(p + qk_\alpha) \times (p + qk_\alpha)$ diagonal matrix. Then the population correlation matrix is defined by

$$\text{(2.3)} \qquad D_{(\alpha)}^{-1/2} \Sigma_{(\alpha)} D_{(\alpha)}^{-1/2} = \begin{bmatrix} P_m & e_{k_\alpha}' \otimes P_{ms} \\ e_{k_\alpha} \otimes P_{ms}' & I_{k_\alpha} \otimes P_s + (e_{k_\alpha} e_{k_\alpha}' - I_{k_\alpha}) \otimes P_{ss} \end{bmatrix}.$$

Srivastava *et al.* (1988) called P_{ms} the interclass correlation matrix and P_{ss} the intraclass correlation matrix. Rao (1945) called the eigenvalues of $P_{ss}P_s^{-1}$ the multivariate intraclass correlations and this will be considered in a later communication.

3. Estimation of interclass and intraclass correlation matrices

Certain classes of estimators for P_{ms} and P_{ss} are introduced to assess the interrelationships among different characteristics. It is convenient to deal with the problem in the canonical form, so we consider the nonsingular transformation suggested by Srivastava (1984). For the α-th family, let

$$\begin{bmatrix} I_p & 0 \\ & \frac{1}{k_\alpha} e'_{k_\alpha} \otimes I_q \\ 0 & A_\alpha \otimes I_q \end{bmatrix} \begin{bmatrix} y_\alpha \\ x_{1\alpha} \\ x_{2\alpha} \\ \vdots \\ x_{k_\alpha,\alpha} \end{bmatrix} = \begin{bmatrix} y_\alpha \\ \bar{x}_\alpha \\ w_{1\alpha} \\ \vdots \\ w_{k_\alpha-1,\alpha} \end{bmatrix}$$

where A_α is a $(k_\alpha - 1) \times k_\alpha$ matrix such that $A_\alpha e_{k_\alpha} = 0$, $A_\alpha A'_\alpha = I_{k_\alpha - 1}$ and $\bar{x}_\alpha = \sum_{j=1}^{k_\alpha} x_{j\alpha}/k_\alpha$. Then it can be seen that

$$\begin{bmatrix} y_\alpha \\ \bar{x}_\alpha \end{bmatrix} \quad \text{and} \quad w_{j\alpha}, \quad j = 1,\ldots,k_\alpha - 1 ,$$

are all uncorrelated, and their mean vectors and covariance matrices are, respectively, given by

$$E\begin{bmatrix} y_\alpha \\ \bar{x}_\alpha \end{bmatrix} = \begin{bmatrix} \mu_m \\ \mu_s \end{bmatrix}, \quad E[w_{j\alpha}] = 0 ,$$

and

$$\operatorname{cov}\left(\begin{bmatrix} y_\alpha \\ \bar{x}_\alpha \end{bmatrix}\right) = \begin{bmatrix} \Sigma_m & \Sigma_{ms} \\ \Sigma'_{ms} & \Sigma_{ss} + \frac{1}{k_\alpha}(\Sigma_s - \Sigma_{ss}) \end{bmatrix}, \quad \operatorname{cov}(w_{j\alpha}) = \Sigma_s - \Sigma_{ss}$$

for $j = 1,\ldots,k_\alpha - 1$.

Let

$$Y = [y_1, y_2, \ldots, y_N], \qquad \bar{X} = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N]$$

and

$$S_\alpha = \sum_{j=1}^{k_\alpha - 1} w_{j\alpha} w'_{j\alpha} = \sum_{j=1}^{k_\alpha} (x_{j\alpha} - \bar{x}_\alpha)(x_{j\alpha} - \bar{x}_\alpha)' .$$

We first construct a class of unbiased estimators of the covariance matrices (2.1) based on Y, $\bar{X}$ and S_α, $\alpha = 1,\ldots,N$. Let B_m and B_s be $N \times N$ positive semidefinite matrices such that $B_m e_N = 0$ and $B_s e_N = 0$ with e_N being the N-dimensional vector of unit elements, and further let B_{ms} be an $N \times N$ matrix such that $B_{ms} e_N = 0$ and $e'_N B_{ms} = 0$. It is assumed that the diagonal elements of B_m, B_s and B_{ms} are of order $O(1)$, and that their off-diagonal elements $O(1/N)$ for the asymptotic distributions of estimators. It may be observed at this stage that Srivastava and Keen (1988) proposed a unified approach for the estimation of univariate interclass correlation coefficient, but the calculation of the order of the asymptotic variances as a function of the weights is left to the reader.

It is easy to verify that

$$
\begin{aligned}
&E[YB_mY'] = (\mathrm{tr}\ B_m)\Sigma_m, \qquad E[YB_{ms}\bar{X}'] = (\mathrm{tr}\ B_{ms})\Sigma_{ms}\,,\\
&E\Big[\bar{X}B_s\bar{X}' + \sum_{\alpha=1}^{N} \omega_\alpha S_\alpha\Big]\\
&\qquad = \Big\{\sum_{\alpha=1}^{N} \omega_\alpha(k_\alpha - 1) + \mathrm{tr}\ B_s D_N^{-1}\Big\}\Sigma_s\\
&\qquad\quad - \Big\{\sum_{\alpha=1}^{N} \omega_\alpha(k_\alpha - 1) - \mathrm{tr}\ B_s(I_N - D_N^{-1})\Big\}\Sigma_{ss}\,,\\
&E\Big[\bar{X}B_s\bar{X}' + \sum_{\alpha=1}^{N} \nu_\alpha S_\alpha\Big]\\
&\qquad = (\mathrm{tr}\ B_s)\Sigma_{ss} + \Big\{\sum_{\alpha=1}^{N} \nu_\alpha(k_\alpha - 1) + \mathrm{tr}\ B_s D_N^{-1}\Big\}(\Sigma_s - \Sigma_{ss})\,,
\end{aligned} \tag{3.1}
$$

where $\omega_1, \omega_2, \ldots, \omega_N$ are non-negative constants, $\nu_1, \nu_2, \ldots, \nu_N$ are constants and $D_N = \mathrm{diag}\,[k_1, k_2, \ldots, k_N]$.

We choose $\{\omega_\alpha;\ \alpha = 1, \ldots, N\}$ and $\{\nu_\alpha;\ \alpha = 1, \ldots, N\}$ such that

$$
\sum_{\alpha=1}^{N} \omega_\alpha(k_\alpha - 1) - \mathrm{tr}\ B_s(I_N - D_N^{-1}) = 0, \qquad \sum_{\alpha=1}^{N} \nu_\alpha(k_\alpha - 1) + \mathrm{tr}\ B_s D_N^{-1} = 0\,. \tag{3.2}
$$

Then, a class of unbiased estimators of Σ_m, Σ_{ms}, Σ_s and Σ_{ss} are, respectively, given by

$$
\begin{aligned}
&\hat{\Sigma}_m = \frac{1}{\mathrm{tr}\ B_m} YB_mY', \qquad \hat{\Sigma}_{ms} = \frac{1}{\mathrm{tr}\ B_{ms}} YB_{ms}\bar{X}'\,,\\
&\hat{\Sigma}_s = \frac{1}{\mathrm{tr}\ B_s}\Big\{\bar{X}B_s\bar{X}' + \sum_{\alpha=1}^{N} \omega_\alpha S_\alpha\Big\},\\
&\hat{\Sigma}_{ss} = \frac{1}{\mathrm{tr}\ B_s}\Big\{\bar{X}B_s\bar{X}' + \sum_{\alpha=1}^{N} \nu_\alpha S_\alpha\Big\}.
\end{aligned} \tag{3.3}
$$

It may be noted that the conditions (3.2) are required to obtain unbiased estimators for Σ_s and Σ_{ss}.

For the diagonal elements of $\hat{\Sigma}_m = (\hat{\sigma}_{ij}^{(m)})$ and $\hat{\Sigma}_s = (\hat{\sigma}_{ij}^{(s)})$, let

$$\hat{D}_{(\alpha)} = \text{diag}\,[\hat{\sigma}_{11}^{(m)},\ldots,\hat{\sigma}_{pp}^{(m)},\hat{\sigma}_{11}^{(s)},\ldots,\hat{\sigma}_{qq}^{(s)},\ldots,\hat{\sigma}_{11}^{(s)},\ldots,\hat{\sigma}_{qq}^{(s)}]\,,$$

a $(p+qk_\alpha)\times(p+qk_\alpha)$ diagonal matrix. The population correlation matrix defined by (2.3) may be estimated by

$$\hat{D}_{(\alpha)}^{-1/2}\hat{\Sigma}_{(\alpha)}\hat{D}_{(\alpha)}^{-1/2} = \begin{bmatrix} \hat{P}_m & e'_{k_\alpha}\otimes\hat{P}_{ms} \\ e_{k_\alpha}\otimes\hat{P}'_{ms} & I_{k_\alpha}\otimes\hat{P}_s + (e_{k_\alpha}e'_{k_\alpha} - I_{k_\alpha})\otimes\hat{P}_{ss} \end{bmatrix}$$

where $\hat{\Sigma}_{(\alpha)}$ is given by replacing Σ_m, Σ_s, Σ_{ms} and Σ_{ss} in (2.2) by their sample estimates (3.3). Hence we have;

$$\hat{P}_{ms} = \text{diag}\,[\hat{\sigma}_{11}^{(m)},\ldots,\hat{\sigma}_{pp}^{(m)}]^{-1/2}\hat{\Sigma}_{ms}\,\text{diag}\,[\hat{\sigma}_{11}^{(s)},\ldots,\hat{\sigma}_{qq}^{(s)}]^{-1/2} \tag{3.4}$$

and

$$\hat{P}_{ss} = \text{diag}\,[\hat{\sigma}_{11}^{(s)},\ldots,\hat{\sigma}_{qq}^{(s)}]^{-1/2}\hat{\Sigma}_{ss}\,\text{diag}\,[\hat{\sigma}_{11}^{(s)},\ldots,\hat{\sigma}_{qq}^{(s)}]^{-1/2}\,, \tag{3.5}$$

where $\text{diag}\,[a_1,\ldots,a_p]^{1/2} = \text{diag}\,[a_1^{1/2},\ldots,a_p^{1/2}]$.

Remark. Suppose we have observations on the concomitant variables for each family. Then, we can eliminate the linear effects due to these variables using the linear model

$$y_\alpha = \mu_m + \beta_1 v_\alpha + \eta_\alpha \quad \text{and} \quad x_{j\alpha} = \mu_s + \beta z_{j\alpha} + a_\alpha + \varepsilon_{j\alpha}; \qquad j = 1,2,\ldots,k_\alpha$$

and $\alpha = 1,2,\ldots,N$, where η_α, a_α and $\varepsilon_{j\alpha}$ are random variables such that

$$E[\eta_\alpha] = 0, \quad E[a_\alpha] = 0, \quad E[\varepsilon_{j\alpha}] = 0\,,$$
$$\text{cov}\,(a_\alpha,\eta_\alpha) = 0, \quad \text{cov}\,(a_\alpha,\varepsilon_{j\alpha}) = 0, \quad \text{cov}\,(\eta_\alpha,\varepsilon_{j\alpha}) = \Sigma_{ms}\,,$$
$$\text{cov}\,(\varepsilon_{j\alpha}) = \Sigma_s - \Sigma_{ss}, \quad \text{cov}\,(a_\alpha) = \Sigma_{ss} \quad \text{and} \quad \text{cov}\,(\eta_\alpha) = \Sigma_m\,.$$

Further, v_α and $z_{j\alpha}$ are observations on concomitant variables on the parent and on the siblings, respectively. Then, we define

$$Y = [y_1,\ldots,y_N], \qquad X_\alpha = [x_{1\alpha},\ldots,x_{k_\alpha,\alpha}], \qquad \bar{X} = [\bar{x}_1,\ldots,\bar{x}_N]\,,$$

$$Z_\alpha = \begin{pmatrix} 1,\ldots,1 \\ z_{1\alpha},\ldots,z_{k_\alpha,\alpha} \end{pmatrix}, \qquad \bar{Z} = \begin{pmatrix} 1,\ldots,1 \\ \bar{z}_1,\ldots,\bar{z}_N \end{pmatrix}, \qquad V = \begin{pmatrix} 1,\ldots,1 \\ v_1,\ldots,v_N \end{pmatrix}$$

where $\bar{x}_\alpha = \sum_{j=1}^{k_\alpha} x_{j\alpha}/k_\alpha$ and $\bar{z}_\alpha = \sum_{j=1}^{k_\alpha} z_{j\alpha}/k_\alpha$. Further, define

$$S_\alpha = X_\alpha Q_{Z_\alpha} X_\alpha' \quad \text{with} \quad Q_{Z_\alpha} = I_{k_\alpha} - Z_\alpha'(Z_\alpha Z_\alpha')^- Z_\alpha \,,$$

where rank $Z_\alpha = r_\alpha$ and rank $Q_{Z_\alpha} = k_\alpha - r_\alpha$ (>0) for $\alpha = 1,\dots, N$. Then we propose the estimates of Σ_m, Σ_s, Σ_{ms} and Σ_{ss} similar to those of (3.3), namely,

$$\hat{\Sigma}_m = \frac{1}{\operatorname{tr} B_m} YB_mY', \qquad \hat{\Sigma}_{ms} = \frac{1}{\operatorname{tr} B_{ms}} YB_{ms}\bar{X}'\,,$$

$$\hat{\Sigma}_s = \frac{1}{\operatorname{tr} B_s}\left\{ \bar{X}B_s\bar{X}' + \sum_{\alpha=1}^{N} \omega_\alpha S_\alpha \right\} \quad \text{and} \quad \hat{\Sigma}_{ss} = \frac{1}{\operatorname{tr} B_s}\left\{ \bar{X}B_s\bar{X}' + \sum_{\alpha=1}^{N} \nu_\alpha S_\alpha \right\}$$

where B_m and B_s are nonnegative definite matrices satisfying the conditions $B_mV' = 0$ (or $VB_m = 0$) and $B_s\bar{Z}' = 0$ (or $\bar{Z}B_s = 0$), and B_{ms} is a matrix such that $B_{ms}\bar{Z}' = 0$ and $VB_{ms} = 0$. For unbiasedness, in place of (3.2), we have

$$\sum_{\alpha=1}^{N} \omega_\alpha(k_\alpha - r_\alpha) - \operatorname{tr} B_s(I_N - D_N^{-1}) = 0 \quad \text{and}$$

$$\sum_{\alpha=1}^{N} \nu_\alpha(k_\alpha - r_\alpha) + \operatorname{tr} B_sD_N^{-1} = 0\,.$$

Here ω_α's are nonnegative and ν_α's are real numbers. For the asymptotic results, we require the similar assumptions as stated for Theorems 4.1 and 4.2. We can define interclass and intraclass correlation matrices as mentioned earlier.

3.1 *Estimates of interclass correlation matrix*

The problem is how to choose the matrices B_m, B_s, B_{ms} and a set of constants $\{\omega_\alpha\}$, $\{\nu_\alpha\}$; $\alpha = 1,\dots, N$ in (3.3). In the univariate situation where $p = q = 1$, various kinds of estimators of the interclass correlation coefficient have been proposed in the literature. Among them, the pairwise, ensemble estimators (Rosner *et al.* (1977)) and Srivastava's (1984) estimator are recommended as stated in the Introduction. These estimators can be generalized quite naturally to the multivariate situation of more than one characteristic, by taking appropriate values for the weights in (3.3).

(1) *Multivariate generalization of the pairwise estimator*

In the univariate situation, the pairwise estimator is obtained by pairing each mother's score with k_α siblings' scores and calculating the ordinary product-moment correlation. This estimator is extended to the one in the multivariate situation of several characteristics.

In (3.3), take

$$B_m = B_s = B_{ms} = D_N - k(N)k'(N) \Big/ \sum_{\alpha=1}^{N} k_\alpha \quad \text{and} \quad \omega_\alpha = 1 , \tag{3.6}$$

where $D_N = \operatorname{diag}[k_1, \ldots, k_N]$ and $k(N) = [k_1, \ldots, k_N]'$ is the N-dimensional vector. Then we have

$$\begin{aligned} \hat{\Sigma}_{m,p} &= (\hat{\sigma}_{ij,p}^{(m)}) = \frac{1}{N_p} \sum_{\alpha=1}^{N} k_\alpha (y_\alpha - \tilde{y})(y_\alpha - \tilde{y})' , \\ \hat{\Sigma}_{s,p} &= (\hat{\sigma}_{ij,p}^{(s)}) = \frac{1}{N_p} \sum_{\alpha=1}^{N} \sum_{j=1}^{k_\alpha} (x_{j\alpha} - \tilde{x})(x_{j\alpha} - \tilde{x})' , \\ \hat{\Sigma}_{ms,p} &= (\hat{\sigma}_{ij,p}^{(ms)}) = \frac{1}{N_p} \sum_{\alpha=1}^{N} \sum_{j=1}^{k_\alpha} (y_\alpha - \tilde{y})(x_{j\alpha} - \tilde{x})' , \end{aligned} \tag{3.7}$$

where $N_p = \sum_{\alpha=1}^{N} \sum_{\beta \neq \alpha}^{N} k_\alpha k_\beta \Big/ \sum_{\alpha=1}^{N} k_\alpha$, $\tilde{y} = \sum_{\alpha=1}^{N} k_\alpha y_\alpha \Big/ \sum_{\alpha=1}^{N} k_\alpha$ and $\tilde{x} = \sum_{\alpha=1}^{N} \sum_{j=1}^{k_\alpha} x_{j\alpha} \Big/ \sum_{\alpha=1}^{N} k_\alpha$. The interclass correlation matrix P_{ms} is estimated by

$$\hat{P}_{ms,p} = \operatorname{diag}[\hat{\sigma}_{11,p}^{(m)}, \ldots, \hat{\sigma}_{pp,p}^{(m)}]^{-1/2} \hat{\Sigma}_{ms,p} \operatorname{diag}[\hat{\sigma}_{11,p}^{(s)}, \ldots, \hat{\sigma}_{qq,p}^{(s)}]^{-1/2} . \tag{3.8}$$

We note that the weights (3.6) do not satisfy the condition in (3.2), so $\hat{\Sigma}_{s,p}$ is not an unbiased estimator for Σ_s. In fact it can be shown that

$$E[\hat{\Sigma}_{s,p}] = \Sigma_s + \left\{ \sum_{\alpha=1}^{N} k_\alpha (k_\alpha - 1) \Big/ \sum_{\alpha=1}^{N} \sum_{\beta \neq \alpha}^{N} k_\alpha k_\beta \right\} (\Sigma_s - \Sigma_{ss}) .$$

An unbiased estimator for Σ_s may be obtained by adjusting the bias term in the above expected value. It can be realized by taking $\omega_\alpha = 1 - k_\alpha \Big/ \sum_{\alpha=1}^{N} k_\alpha$ in (3.3), and then we have an unbiased estimator for Σ_s in the form

$$\hat{\Sigma}_{s,pc} = (\hat{\sigma}_{ij,pc}^{(s)}) = \hat{\Sigma}_{s,p} - \sum_{\alpha=1}^{N} k_\alpha S_\alpha \Big/ \left(N_p \sum_{\alpha=1}^{N} k_\alpha \right) .$$

An estimator of P_{ms} is given by

$$\hat{P}_{ms,pc} = \operatorname{diag}[\hat{\sigma}_{11,p}^{(m)}, \ldots, \hat{\sigma}_{pp,p}^{(m)}]^{-1/2} \hat{\Sigma}_{ms,p} \operatorname{diag}[\hat{\sigma}_{11,pc}^{(s)}, \ldots, \hat{\sigma}_{qq,pc}^{(s)}]^{-1/2} . \tag{3.9}$$

(2) *Multivariate generalization of the ensemble estimator*

A multivariate analogue of the ensemble estimator is obtained by taking

$$B_m = B_s = B_{ms} = I_N - \frac{1}{N} e_N e_N' \quad \text{and} \quad \omega_\alpha = \frac{N-1}{N k_\alpha}$$

in (3.3). It can be readily checked that these weights satisfy the condition in (3.2), so that unbiased estimators for Σ_m, Σ_s and Σ_{ms} are, respectively, given by

$$\hat{\Sigma}_{m,e} = (\hat{\sigma}_{ij,e}^{(m)}) = \frac{1}{N-1} \sum_{\alpha=1}^{N} (y_\alpha - \bar{y})(y_\alpha - \bar{y})' ,$$

$$(3.10) \qquad \hat{\Sigma}_{s,e} = (\hat{\sigma}_{ij,e}^{(s)}) = \frac{1}{N-1} \left\{ \sum_{\alpha=1}^{N} (\bar{x}_\alpha - \bar{x})(\bar{x}_\alpha - \bar{x})' + \frac{N-1}{N} \sum_{\alpha=1}^{N} \frac{1}{k_\alpha} S_\alpha \right\} ,$$

$$\hat{\Sigma}_{ms,e} = (\hat{\sigma}_{ij,e}^{(ms)}) = \frac{1}{N-1} \sum_{\alpha=1}^{N} (y_\alpha - \bar{y})(\bar{x}_\alpha - \bar{x})' ,$$

where $\bar{y} = \sum_{\alpha=1}^{N} y_\alpha / N$, $\bar{x}_\alpha = \sum_{j=1}^{k_\alpha} x_{j\alpha} / k_\alpha$ and $\bar{x} = \sum_{\alpha=1}^{N} \bar{x}_\alpha / N$. Then an estimator of P_{ms} based on (3.10) is of the form

$$(3.11) \qquad \hat{P}_{ms,e} = \operatorname{diag}\,[\hat{\sigma}_{11,e}^{(m)}, \ldots, \hat{\sigma}_{pp,e}^{(m)}]^{-1/2} \hat{\Sigma}_{ms,e} \operatorname{diag}\,[\hat{\sigma}_{11,e}^{(s)}, \ldots, \hat{\sigma}_{qq,e}^{(s)}]^{-1/2} .$$

(3) *Multivariate generalization of Srivastava's estimator*
To obtain the estimator given by Srivastava *et al.* (1988), set

$$(3.12) \qquad \begin{aligned} &B_m = B_s = B_{ms} = I_N - \frac{1}{N} e_N e_N' \quad \text{and} \\ &\omega_\alpha = \frac{N-1}{N} \sum_{\alpha=1}^{N} \frac{1}{k_\alpha} (k_\alpha - 1) \Big/ \sum_{\alpha=1}^{N} (k_\alpha - 1) = \omega_s, \quad \text{say} , \end{aligned}$$

which satisfy the condition in (3.2). Unbiased estimators for Σ_m, Σ_s and Σ_{ms} are, respectively, given by

$$\hat{\Sigma}_{m,s} = (\hat{\sigma}_{ij,s}^{(m)}) = \frac{1}{N-1} \sum_{\alpha=1}^{N} (y_\alpha - \bar{y})(y_\alpha - \bar{y})' ,$$

$$(3.13) \qquad \hat{\Sigma}_{s,s} = (\hat{\sigma}_{ij,s}^{(s)}) = \frac{1}{N-1} \left\{ \sum_{\alpha=1}^{N} (\bar{x}_\alpha - \bar{x})(\bar{x}_\alpha - \bar{x})' + \omega_s \sum_{\alpha=1}^{N} S_\alpha \right\} ,$$

$$\hat{\Sigma}_{ms,s} = (\hat{\sigma}_{ij,s}^{(ms)}) = \frac{1}{N-1} \sum_{\alpha=1}^{N} (y_\alpha - \bar{y})(\bar{x}_\alpha - \bar{x})' ,$$

where $\bar{y}$, $\bar{x}_\alpha$ and $\bar{x}$ are given in (3.10). Hence an estimator of P_{ms} based on (3.13) is

$$(3.14) \qquad \hat{P}_{ms,s} = \operatorname{diag}\,[\hat{\sigma}_{11,s}^{(m)}, \ldots, \hat{\sigma}_{pp,s}^{(m)}]^{-1/2} \hat{\Sigma}_{ms,s} \operatorname{diag}\,[\hat{\sigma}_{11,s}^{(s)}, \ldots, \hat{\sigma}_{qq,s}^{(s)}]^{-1/2} .$$

The asymptotic variance of each element of this estimator was given by Srivastava *et al.* (1988).

(4) In the univariate case where $p = q = 1$, Srivastava and Keen's (1988) general estimator can be obtained by taking

$$B_m = D_v - \frac{vv'}{V}, \quad B_s = D_u - \frac{uu'}{U}, \quad B_{ms} = D_t - \frac{vt'}{V} - \frac{tu'}{U} + T\frac{vu'}{UV},$$

$$\omega_\alpha = \mathrm{tr}\, B_s(I_N - D_N^{-1}) \Big/ \sum_{\alpha=1}^{N} (k_\alpha - 1) \quad \text{for} \quad \alpha = 1,\dots, N,$$

where $t' = (t_1,\dots, t_N)$, $u' = (u_1,\dots, u_N)$, $v' = (v_1,\dots, v_N)$, $D_t = \mathrm{diag}\,[t_1,\dots, t_N]$, $D_u = \mathrm{diag}\,[u_1,\dots, u_N]$, $D_v = \mathrm{diag}\,[v_1,\dots, v_N]$, $T = \sum_{\alpha=1}^{N} t_\alpha$, $U = \sum_{\alpha=1}^{N} u_\alpha$ and $V = \sum_{\alpha=1}^{N} v_\alpha$. For the asymptotic distribution, we shall require conditions similar to those in Theorem 4.1. The corrected pairwise estimator proposed by Srivastava and Keen ((1988), p. 733) can be generalized to the multivariate situation of several characteristics, using the above relations. In (3.3), taking (3.6) for B_m, B_s, B_{ms} and

$$\omega_\alpha = \left(\sum_{\alpha=1}^{N} k_\alpha - \sum_{\alpha=1}^{N} k_\alpha^2 \Big/ \sum_{\alpha=1}^{N} k_\alpha - N + 1 \right) \Big/ \sum_{\alpha=1}^{N} (k_\alpha - 1) = \omega_{ps}, \quad \text{say},$$

we have the unbiased estimators $\hat{\Sigma}_{m,p}$, $\hat{\Sigma}_{ms,p}$ in (3.7) and

$$\hat{\Sigma}_{s,ps} = (\hat{\sigma}_{ij,ps}^{(s)}) = \frac{1}{N_p} \left\{ \sum_{\alpha=1}^{N} k_\alpha (\bar{x}_\alpha - \tilde{x})(\bar{x}_\alpha - \tilde{x})' + \omega_{ps} \sum_{\alpha=1}^{N} S_\alpha \right\}.$$

Then an estimate of P_{ms} is given by

$$\text{(3.15)} \quad \hat{P}_{ms,ps} = \mathrm{diag}\,[\hat{\sigma}_{11,p}^{(m)},\dots, \hat{\sigma}_{pp,p}^{(m)}]^{-1/2} \hat{\Sigma}_{ms,p}\, \mathrm{diag}\,[\hat{\sigma}_{11,ps}^{(s)},\dots, \hat{\sigma}_{qq,ps}^{(s)}]^{-1/2}.$$

3.2 *Estimates of intraclass correlation matrix*

(1) *Multivariate generalization of the analysis of variance estimator*

A multivariate analogue of the analysis of variance estimator (Donner and Koval (1980)) can be obtained by taking, in (3.3),

$$\text{(3.16)} \quad \begin{aligned} B_s &= D_N - k(N)k(N)' \Big/ \sum_{\alpha=1}^{N} k_\alpha, \\ \omega_\alpha &= (N_p - N + 1) \Big/ \sum_{\alpha=1}^{N} (k_\alpha - 1) = \omega_a, \quad \text{say}, \\ \nu_\alpha &= -(N - 1) \Big/ \sum_{\alpha=1}^{N} (k_\alpha - 1) = \nu_a, \quad \text{say}, \end{aligned}$$

for $\alpha = 1,\ldots, N$, where D_N, $k(N)$ and N_p are defined in (1) of Subsection 3.1. The unbiased estimators of Σ_s and Σ_{ss} are, respectively, given by

$$\hat{\Sigma}_{s,a} = (\hat{\sigma}^{(s)}_{ij,a}) = \frac{1}{N_p}\left\{\sum_{\alpha=1}^{N} k_\alpha(\bar{x}_\alpha - \tilde{x})(\bar{x}_\alpha - \tilde{x})' + \omega_a \sum_{\alpha=1}^{N} S_\alpha\right\},$$

$$\hat{\Sigma}_{ss,a} = (\hat{\sigma}^{(ss)}_{ij,a}) = \frac{1}{N_p}\left\{\sum_{\alpha=1}^{N} k_\alpha(\bar{x}_\alpha - \tilde{x})(\bar{x}_\alpha - \tilde{x})' + \nu_a \sum_{\alpha=1}^{N} S_\alpha\right\},$$

where $\tilde{x}$ is defined in (3.7). These estimators were originally introduced by Rao (1953). Then we have an estimator of the intraclass correlation matrix P_{ss} in the form

$$\hat{P}_{ss,a} = \mathrm{diag}\,[\hat{\sigma}^{(s)}_{11,a},\ldots, \hat{\sigma}^{(s)}_{qq,a}]^{-1/2}\hat{\Sigma}_{ss,a}\,\mathrm{diag}\,[\hat{\sigma}^{(s)}_{11,a},\ldots, \hat{\sigma}^{(s)}_{qq,a}]^{-1/2}. \tag{3.17}$$

(2) *Multivariate generalization of a weighted pairwise estimator*

A weighted pairwise intraclass correlation estimator proposed by Karlin *et al.* (1981) is extended by setting, in (3.3),

$$\begin{aligned} &B_s = \mathrm{diag}\,[u_1, u_2,\ldots, u_N] - uu'/U\,, \\ &\omega_\alpha = u_\alpha/k_\alpha, \qquad \nu_\alpha = -\,u_\alpha/\{k_\alpha(k_\alpha - 1)\}, \qquad \alpha = 1,\ldots, N \end{aligned} \tag{3.18}$$

where $u' = (u_1,\ldots, u_N)$ and $\sum_{\alpha=1}^{N} u_\alpha = U$. It is assumed that u_α's are of order $O(1)$ as N tends to infinity. Then it can be seen that

$$\hat{\Sigma}_{s,k} = (\hat{\sigma}^{(s)}_{ij,k}) = \frac{1}{N_u}\sum_{\alpha=1}^{N}\frac{u_\alpha}{k_\alpha}\sum_{j=1}^{k_\alpha}(x_{j\alpha} - \hat{\mu})(x_{j\alpha} - \hat{\mu})',$$

$$\hat{\Sigma}_{ss,k} = (\hat{\sigma}^{(ss)}_{ij,k}) = \frac{1}{N_u}\sum_{\alpha=1}^{N}\frac{u_\alpha}{k_\alpha(k_\alpha - 1)}\sum_{i=1}^{k_\alpha}\sum_{j\neq i}^{k_\alpha}(x_{i\alpha} - \hat{\mu})(x_{j\alpha} - \hat{\mu})',$$

where $\hat{\mu} = (1/U)\sum_{\alpha=1}^{N}(u_\alpha/k_\alpha)\sum_{j=1}^{k_\alpha} x_{j\alpha}$ and $N_u = U - \sum_{\alpha=1}^{N} u_\alpha^2/U$. It follows from (3.1) that

$$E[\hat{\Sigma}_{s,k}] = \Sigma_s + \frac{1}{N_u U}\sum_{\alpha=1}^{N}\frac{1}{k_\alpha}(k_\alpha - 1)u_\alpha^2(\Sigma_s - \Sigma_{ss}),$$

$$E[\hat{\Sigma}_{ss,k}] = \Sigma_{ss} - \frac{1}{N_u U}\sum_{\alpha=1}^{N}\frac{1}{k_\alpha}u_\alpha^2(\Sigma_s - \Sigma_{ss}).$$

Since $(1/U^2)\sum_\alpha u_\alpha^2(k_\alpha - 1)/k_\alpha$ and $(1/U^2)\sum_\alpha u_\alpha^2/k_\alpha$ converge to 0 as N tends to infinity, $\hat{\Sigma}_{s,k}$ and $\hat{\Sigma}_{ss,k}$ are consistent estimators for Σ_s and Σ_{ss}, respectively. Then we have

$$(3.19)\qquad \hat{P}_{ss,k} = \operatorname{diag}[\hat{\sigma}^{(s)}_{11,k},\ldots,\hat{\sigma}^{(s)}_{qq,k}]^{-1/2}\hat{\Sigma}_{ss,k}\operatorname{diag}[\hat{\sigma}^{(s)}_{11,k},\ldots,\hat{\sigma}^{(s)}_{qq,k}]^{-1/2}.$$

It might be noted that a multivariate extension of the pairwise intraclass correlation estimator is obtained by putting $u_\alpha = k_\alpha(k_\alpha - 1)$.

Srivastava *et al.* (1988) proposed an estimator of the intraclass correlation matrix P_{ss}, which can be obtained by (3.12) and setting

$$v_\alpha = -\frac{N-1}{N}\sum_{\alpha=1}^{N}\frac{1}{k_\alpha}\bigg/\sum_{\alpha=1}^{N}(k_\alpha - 1)$$

in (3.3). They also derived the asymptotic variance of the element of their proposed estimator $\hat{P}_{ss,s}$. It can be easily seen that the weights satisfy the condition (3.2).

4. Inference procedures for the interclass and intraclass correlations

4.1 *Asymptotic results*

Under the assumption of multivariate normality we first derive asymptotic distributions of the estimators $\hat{P}_{ms}$ and $\hat{P}_{ss}$ defined by (3.4) and (3.5), respectively. The results are used to obtain asymptotic distributions of several estimators introduced in the last section.

Suppose that $z_{(\alpha)}$ $\alpha = 1,\ldots,N$, are independently and normally distributed with covariance matrix (2.2). Let $B_m = (b^{(m)}_{\alpha\beta})$, $B_s = (b^{(s)}_{\alpha\beta})$ and $B_{ms} = (b^{(ms)}_{\alpha\beta})$. It is assumed that for $\alpha = 1,\ldots,N$

$$b^{(m)}_{\alpha\alpha} \to b^{(m)}_{\alpha}, \qquad b^{(s)}_{\alpha\alpha} \to b^{(s)}_{\alpha} \qquad \text{and} \qquad b^{(ms)}_{\alpha\alpha} \to b^{(ms)}_{\alpha}$$

as N tends to infinity. If the weights ω_α and v_α ($\alpha = 1,\ldots,N$) depend on the sample size N, ω_α and v_α are regarded as their limiting values ($< \infty$). It is further assumed that the off-diagonal elements of B_m, B_s and B_{ms} are of order $O(1/N)$. Then, the asymptotic results are summarized in the following Theorems 4.1 and 4.2.

THEOREM 4.1. *Let $\hat{P}_{ms} = (\hat{\rho}^{(ms)}_{ij})$ be the estimator, defined by* (3.4), *of the interclass correlation matrix $P_{ms} = (\rho^{(ms)}_{ij})$. Then, $\sqrt{N}(\hat{P}_{ms} - P_{ms})$ is asymptotically normally distributed with mean* 0 *and covariance matrix having elements*

$$(4.1)\qquad \operatorname{var}\{\sqrt{N}(\hat{\rho}^{(ms)}_{ij} - \rho^{(ms)}_{ij})\}$$

$$= N\bigg[\frac{1}{N_m N_s}\sum_{\alpha=1}^{N} b^{(m)}_{\alpha} b^{(s)}_{\alpha} \rho^{(ms)^4}_{ij}$$

$$+\left\{\frac{1}{2N_m^2}\sum_{\alpha=1}^{N} b_\alpha^{(m)^2} - \frac{2}{N_m N_{ms}}\sum_{\alpha=1}^{N} b_\alpha^{(m)} b_\alpha^{(ms)} + \frac{1}{N_{ms}^2}\sum_{\alpha=1}^{N} b_\alpha^{(ms)^2}\right.$$

$$\left.+\frac{1}{2N_s^2}\sum_{\alpha=1}^{N} b_\alpha^{(s)^2} d_{jj,\alpha}^2 - \frac{2}{N_s N_{ms}}\sum_{\alpha=1}^{N} b_\alpha^{(s)} b_\alpha^{(ms)} d_{jj,\alpha}\right\}\rho_{ij}^{(ms)^2}$$

$$+\frac{1}{N_{ms}^2}\sum_{\alpha=1}^{N} b_\alpha^{(ms)^2} d_{jj,\alpha} + \frac{1}{2N_s^2}\varphi_{jj}^2 \rho_{ij}^{(ms)^2}\sum_{\alpha=1}^{N}(k_\alpha - 1)\omega_\alpha^2\Bigg],$$

$$\operatorname{cov}\{\sqrt{N}(\hat{\rho}_{ij}^{(ms)} - \rho_{ij}^{(ms)}), \sqrt{N}(\hat{\rho}_{kl}^{(ms)} - \rho_{kl}^{(ms)})\}$$

$$= N\Bigg[\frac{1}{2}\left\{\frac{1}{N_m N_s}\sum_{\alpha=1}^{N} b_\alpha^{(m)} b_\alpha^{(s)}(\rho_{il}^{(ms)^2} + \rho_{kj}^{(ms)^2})\right.$$

$$\left.+\frac{1}{N_m^2}\sum_{\alpha=1}^{N} b_\alpha^{(m)^2}\rho_{ik}^{(m)^2} + \frac{1}{N_s^2}\sum_{\alpha=1}^{N} b_\alpha^{(s)^2} d_{jl,\alpha}^2\right\}\rho_{ij}^{(ms)}\rho_{kl}^{(ms)}$$

$$-\frac{1}{N_m N_{ms}}\sum_{\alpha=1}^{N} b_\alpha^{(m)} b_\alpha^{(ms)}\rho_{ik}^{(m)}(\rho_{ij}^{(ms)}\rho_{il}^{(ms)} + \rho_{kj}^{(ms)}\rho_{kl}^{(ms)})$$

$$-\frac{1}{N_s N_{ms}}\sum_{\alpha=1}^{N} b_\alpha^{(s)} b_\alpha^{(ms)} d_{jl,\alpha}(\rho_{ij}^{(ms)}\rho_{kj}^{(ms)} + \rho_{il}^{(ms)}\rho_{kl}^{(ms)})$$

$$+\frac{1}{N_{ms}^2}\sum_{\alpha=1}^{N} b_\alpha^{(ms)^2}(\rho_{il}^{(ms)}\rho_{kj}^{(ms)} + d_{jl,\alpha}\rho_{ik}^{(m)})$$

$$+\frac{1}{2N_s^2}\varphi_{jl}^2\rho_{ij}^{(ms)}\rho_{kl}^{(ms)}\sum_{\alpha=1}^{N}(k_\alpha - 1)\omega_\alpha^2\Bigg],$$

where $N_m = \sum_{\alpha=1}^{N} b_\alpha^{(m)}$, $N_s = \sum_{\alpha=1}^{N} b_\alpha^{(s)}$, $N_{ms} = \sum_{\alpha=1}^{N} b_\alpha^{(ms)}$, $P_m = (\rho_{ij}^{(m)})$, $P_s = (\rho_{ij}^{(s)})$ *and*

$$d_{jj,\alpha} = \{1 + (k_\alpha - 1)\rho_{jj}^{(ss)}\}/k_\alpha, \qquad \varphi_{jj} = 1 - \rho_{jj}^{(ss)},$$
$$d_{jl,\alpha} = \{\rho_{jl}^{(s)} + (k_\alpha - 1)\rho_{jl}^{(ss)}\}/k_\alpha, \qquad \varphi_{jl} = \rho_{jl}^{(s)} - \rho_{jl}^{(ss)}. \tag{4.2}$$

The bias of $\hat{\rho}_{ij}^{(ms)}$ *is asymptotically given by*

$$E[\hat{\rho}_{ij}^{(ms)} - \rho_{ij}^{(ms)}] \approx \rho_{ij}^{(ms)}\Bigg[\frac{1}{2N_m N_s}\sum_{\alpha=1}^{N} b_\alpha^{(m)} b_\alpha^{(s)}\rho_{ij}^{(ms)^2}$$

$$+\frac{3}{4N_m^2}\sum_{\alpha=1}^{N} b_\alpha^{(m)^2} - \frac{1}{N_m N_{ms}}\sum_{\alpha=1}^{N} b_\alpha^{(m)} b_\alpha^{(ms)}$$

$$-\frac{1}{N_s N_{ms}}\sum_{\alpha=1}^{N} b_\alpha^{(s)} b_\alpha^{(ms)} d_{jj,\alpha}$$

$$+\frac{3}{4N_s^2}\sum_{\alpha=1}^{N} b_\alpha^{(s)^2} d_{jj,\alpha}^2 + \frac{3}{4N_s^2}\varphi_{jj}^2\sum_{\alpha=1}^{N}(k_\alpha - 1)\omega_\alpha^2\Bigg].$$

By analogy with the univariate approach, we proposed several estimators of the interclass correlation matrix P_{ms}. We recall that these estimators were obtained by taking equal weights for the matrices B_m, B_s and B_{ms} in (3.3) except Srivastava and Keen's general estimator given in (4) of Subsection 3.1. The following corollary gives asymptotic results for the case of $B_m = B_s = B_{ms}$.

COROLLARY 4.1. *Suppose* $B_m = B_s = B_{ms} = B$, *say. It is assumed that for* $\alpha = 1,\dots, N$ *the* α*-th diagonal element of* B *converges to* b_α *as* N *tends to infinity. Then the asymptotic distribution of* $\sqrt{N}(\hat{\rho}_{ij}^{(ms)} - \rho_{ij}^{(ms)})$ *is normal with mean* 0 *and variance*

$$\tau_{ij}^2 = N\left(\sum_{\alpha=1}^{N} b_\alpha\right)^{-2}\left\{\sum_{\alpha=1}^{N} b_\alpha^2 a_{ij,\alpha} + \frac{1}{2}(1-\rho_{jj}^{(ss)})^2\rho_{ij}^{(ms)^2}\sum_{\alpha=1}^{N}(k_\alpha - 1)\omega_\alpha^2\right\}, \tag{4.3}$$

where

$$a_{ij,\alpha} = \rho_{ij}^{(ms)^4} + \frac{1}{2}(d_{jj,\alpha}^2 - 4d_{jj,\alpha} - 1)\rho_{ij}^{(ms)^2} + d_{jj,\alpha} \tag{4.4}$$

with $d_{jj,\alpha} = \{1 + (k_\alpha - 1)\rho_{jj}^{(ss)}\}/k_\alpha$.

The asymptotic bias of $\hat{\rho}_{ij}^{(ms)}$ *is*

$$E[\hat{\rho}_{ij}^{(ms)} - \rho_{ij}^{(ms)}] \approx \left(\sum_{\alpha=1}^{N} b_\alpha\right)^{-2}\rho_{ij}^{(ms)} \cdot\left\{\sum_{\alpha=1}^{N} b_\alpha^2 c_{ij,\alpha} + \frac{3}{4}(1-\rho_{jj}^{(ss)})^2\sum_{\alpha=1}^{N}(k_\alpha - 1)\omega_\alpha^2\right\}, \tag{4.5}$$

where

$$c_{ij,\alpha} = \frac{1}{2}\rho_{ij}^{(ms)^2} + \frac{1}{4}d_{jj,\alpha}(3d_{jj,\alpha} - 4) - \frac{1}{4}. \tag{4.6}$$

THEOREM 4.2. *Let* $\hat{P}_{ss} = (\hat{\rho}_{ij}^{(ss)})$ *be the estimator, defined by* (3.5), *of the intraclass correlation matrix* $P_{ss} = (\rho_{ij}^{(ss)})$. *Then* $\sqrt{N}(\hat{P}_{ss} - P_{ss})$ *is asymptotically normally distributed with mean* 0 *and covariance matrix having elements*

$$\operatorname{var}\{\sqrt{N}(\hat{\rho}_{ij}^{(ss)} - \rho_{ij}^{(ss)})\} = \frac{N}{N_s^2}\left[\sum_{\alpha=1}^{N} b_\alpha^{(s)^2}\left\{d_{ii,\alpha}d_{jj,\alpha} + d_{ij,\alpha}^2 - 2\rho_{ij}^{(ss)}d_{ij,\alpha}(d_{ii,\alpha} + d_{jj,\alpha})\right.\right. \tag{4.7}$$

$$+\frac{1}{2}\rho_{ij}^{(ss)^2}(d_{ii,\alpha}^2+d_{jj,\alpha}^2+2d_{ij,\alpha}^2)\Big\}$$

$$+\sum_{\alpha=1}^{N}(k_\alpha-1)\Big\{v_\alpha^2(\varphi_{ii}\varphi_{jj}+\varphi_{ij}^2)-2\omega_\alpha v_\alpha\rho_{ij}^{(ss)}\varphi_{ij}(\varphi_{ii}+\varphi_{jj})$$

$$+\frac{1}{2}\omega_\alpha^2\rho_{ij}^{(ss)^2}(\varphi_{ii}^2+\varphi_{jj}^2+2\varphi_{ij}^2)\Big\}\Big],$$

$$\mathrm{cov}\,\{\sqrt{N}(\hat\rho_{ij}^{(ss)}-\rho_{ij}^{(ss)}),\sqrt{N}(\hat\rho_{kl}^{(ss)}-\rho_{kl}^{(ss)})\}$$

$$=\frac{N}{N_s^2}\Big[\sum_{\alpha=1}^{N}b_\alpha^{(s)^2}\Big\{d_{ik,\alpha}d_{jl,\alpha}+d_{il,\alpha}d_{jk,\alpha}$$

$$-\rho_{ij}^{(ss)}(d_{ik,\alpha}d_{il,\alpha}+d_{jk,\alpha}d_{jl,\alpha})-\rho_{kl}^{(ss)}(d_{ik,\alpha}d_{jk,\alpha}+d_{il,\alpha}d_{jl,\alpha})$$

$$+\frac{1}{2}\rho_{ij}^{(ss)}\rho_{kl}^{(ss)}(d_{ik,\alpha}^2+d_{il,\alpha}^2+d_{jk,\alpha}^2+d_{jl,\alpha}^2)\Big\}$$

$$+\sum_{\alpha=1}^{N}(k_\alpha-1)\Big\{v_\alpha^2(\varphi_{ik}\varphi_{jl}+\varphi_{il}\varphi_{jk})$$

$$-\omega_\alpha v_\alpha\rho_{ij}^{(ss)}(\varphi_{ik}\varphi_{il}+\varphi_{jk}\varphi_{jl})-\omega_\alpha v_\alpha\rho_{kl}^{(ss)}(\varphi_{ik}\varphi_{jk}+\varphi_{il}\varphi_{jl})$$

$$+\frac{1}{2}\omega_\alpha^2\rho_{ij}^{(ss)}\rho_{kl}^{(ss)}(\varphi_{ik}^2+\varphi_{il}^2+\varphi_{jk}^2+\varphi_{jl}^2)\Big\}\Big],$$

where $N_s=\sum_{\alpha=1}^{N}b_\alpha^{(s)}$ *and* $d_{ij,\alpha}$ *and* φ_{ij} *are defined in* (4.2).

The asymptotic bias of $\hat\rho_{ij}^{(ss)}$ *is*

$$E[\hat\rho_{ij}^{(ss)}-\rho_{ij}^{(ss)}]\approx\frac{1}{N_s^2}\rho_{ij}^{(ss)}\Big[\sum_{\alpha=1}^{N}b_\alpha^{(s)^2}\Big\{\frac{3}{4}(d_{ii,\alpha}^2+d_{jj,\alpha}^2)$$

$$+\frac{1}{2}d_{ij,\alpha}^2-\rho_{ij}^{(ss)^{-1}}d_{ij,\alpha}(d_{ii,\alpha}+d_{jj,\alpha})\Big\}$$

$$+\Big\{\frac{3}{4}(\varphi_{ii}^2+\varphi_{jj}^2)+\frac{1}{2}\varphi_{ij}^2\Big\}\sum_{\alpha=1}^{N}(k_\alpha-1)\omega_\alpha^2$$

$$-\rho_{ij}^{(ss)^{-1}}\varphi_{ij}(\varphi_{ii}+\varphi_{jj})\sum_{\alpha=1}^{N}(k_\alpha-1)\omega_\alpha v_\alpha\Big].$$

The asymptotic results in Theorems 4.1 and 4.2 can be obtained, using an approach discussed in Konishi ((1982), p. 513). Theorems 4.1 and 4.2 present unified formulae to asymptotic distributions of estimators of P_{ms} and P_{ss}, constructed by the unbiased estimators (3.3). It might be, however, noted that these theorems can also be used to derive asymptotic distributions of estimators based on consistent estimators of the covariance matrices instead of the unbiased estimators (3.3), but slight modification is required for deriving asymptotic biases. This will be discussed through the case of the multivariate extension of the pairwise estimator in the following example.

Example 1. The estimators $\hat{P}_{ms,p}$ in (3.8), $\hat{P}_{ms,pc}$ in (3.9) and $\hat{P}_{ms,ps}$ in (3.15) may be regarded as a multivariate generalization of the pairwise estimator. The $\hat{P}_{ms,pc}$ and $\hat{P}_{ms,ps}$ were constructed based on the unbiased estimators (3.3) for which the weights satisfy the condition in (3.2), so Corollary 4.1 can be directly applied to derive their asymptotic distributions. On the other hand, the weights for $\hat{P}_{ms,p}$ do not satisfy the condition in (3.2), since $\hat{\Sigma}_{s,p}$ is not unbiased for Σ_s. It should be, however, noted that the Taylor series expansion of $\hat{P}_{ms,p}$ around P_{ms} agrees with that of $\hat{P}_{ms,pc}$ to order $n^{-1/2}$. This indicates that $\hat{P}_{ms,p}$ and $\hat{P}_{ms,pc}$ have the same limiting distribution, although the asymptotic bias of $\hat{P}_{ms,p}$ differs from that of $\hat{P}_{ms,pc}$.

We recall that $\hat{P}_{ms,pc}$ was obtained by taking $b_{\alpha\alpha}^{(m)} = b_{\alpha\alpha}^{(s)} = b_{\alpha\alpha}^{(ms)} = k_\alpha - k_\alpha^2 \Big/ \sum_\alpha k_\alpha$ and $\omega_\alpha = 1 - k_\alpha \Big/ \sum_\alpha k_\alpha$ for $\alpha = 1,\ldots, N$ in (3.3). Hence, put $b_\alpha = k_\alpha$ and $\omega_\alpha = 1$ ($\alpha = 1,\ldots, N$) in Corollary 4.1. Then, for $\hat{P}_{ms,pc} = (\hat{\rho}_{ij,pc}^{(ms)})$, $\sqrt{N}(\hat{\rho}_{ij,pc}^{(ms)} - \rho_{ij}^{(ms)})$ is asymptotically normally distributed with mean 0 and variance

$$\tau_{ij,pc}^2 = N\left(\sum_{\alpha=1}^N k_\alpha\right)^{-2}\left\{\sum_{\alpha=1}^N k_\alpha^2 a_{ij,\alpha} + \frac{1}{2}\rho_{ij}^{(ms)^2}(1-\rho_{jj}^{(ss)})^2\sum_{\alpha=1}^N (k_\alpha - 1)\right\}$$

where $a_{ij,\alpha}$ is given by (4.4). For $\hat{P}_{ms,p} = (\hat{\rho}_{ij,p}^{(ms)})$ or $\hat{P}_{ms,ps} = (\hat{\rho}_{ij,ps}^{(ms)})$, the asymptotic distribution of $\hat{\rho}_{ij,p}^{(ms)}$ or $\hat{\rho}_{ij,ps}^{(ms)}$ is the same as that of $\hat{\rho}_{ij,pc}^{(ms)}$.

The asymptotic biases of $\hat{\rho}_{ij,pc}^{(ms)}$ and $\hat{\rho}_{ij,p}^{(ms)}$ are, respectively, given by

$$\mathrm{ab}(\hat{\rho}_{ij,pc}^{(ms)}) = \left(\sum_{\alpha=1}^N k_\alpha\right)^{-2}\rho_{ij}^{(ms)}\left\{\sum_{\alpha=1}^N k_\alpha^2 c_{ij,\alpha} + \frac{3}{4}(1-\rho_{jj}^{(ss)})^2\sum_{\alpha=1}^N (k_\alpha - 1)\right\},$$

$$\mathrm{ab}(\hat{\rho}_{ij,p}^{(ms)}) = ab(\hat{\rho}_{ij,pc}^{(ms)}) - \frac{1}{2}\rho_{ij}^{(ms)}(1-\rho_{jj}^{(ss)})\sum_{\alpha=1}^N k_\alpha(k_\alpha - 1) \Bigg/ \left(\sum_{\alpha=1}^N k_\alpha\right)^2$$

where $c_{ij,\alpha}$ is given by (4.6).

Example 2. Let $\hat{P}_{ms,e} = (\hat{\rho}_{ij,e}^{(ms)})$ and $\hat{P}_{ms,s} = (\hat{\rho}_{ij,s}^{(ms)})$ be the estimators defined by (3.11) and (3.14), respectively. These estimators were constructed based on the unbiased estimators (3.3) with the same weights $B_m = B_s = B_{ms} = I_n - e_N e_N'/N$. Hence asymptotic results can be represented in a common formula for both estimators. Let $\hat{\rho}_{ij,\cdot}^{(ms)}$ denote $\hat{\rho}_{ij,e}^{(ms)}$ or $\hat{\rho}_{ij,s}^{(ms)}$. Then it follows from Corollary 4.1 that $\sqrt{N}(\hat{\rho}_{ij,\cdot}^{(ms)} - \rho_{ij}^{(ms)})$ is asymptotically normally distributed with mean 0 and variance

$$\tau_{ij,\cdot}^2 = \frac{1}{N}\left\{ \sum_{\alpha=1}^{N} a_{ij,\alpha} + \frac{1}{2}\rho_{ij}^{(ms)^2}(1-\rho_{jj}^{(ss)})^2\omega_{\cdot} \right\}$$

where $a_{ij,\alpha}$ is given by (4.4) and corresponding to each of $\hat{\rho}_{ij,e}^{(ms)}$ or $\hat{\rho}_{ij,s}^{(ms)}$, $\omega_{\cdot}$ is represented by

$$\text{(4.8)}\qquad \begin{aligned} \omega_e &= \sum_{\alpha=1}^{N}(k_\alpha - 1)/k_\alpha^2 && \text{for } \hat{\rho}_{ij,e}^{(ms)}, \\ \omega_s &= \left\{\sum_{\alpha=1}^{N}(k_\alpha-1)/k_\alpha\right\}^2 \Big/ \sum_{\alpha=1}^{N}(k_\alpha-1) && \text{for } \hat{\rho}_{ij,s}^{(ms)}. \end{aligned}$$

The asymptotic bias of $\hat{\rho}_{ij,\cdot}^{(ms)}$ is

$$\text{ab}(\hat{\rho}_{ij,\cdot}^{(ms)}) = \frac{1}{N^2}\rho_{ij}^{(ms)}\left\{ \sum_{\alpha=1}^{N} c_{ij,\alpha} + \frac{3}{4}(1-\rho_{jj}^{(ss)})^2\omega_{\cdot} \right\}$$

where $c_{ij,\alpha}$ and $\omega_{\cdot}$ are, respectively, given by (4.6) and (4.8).

Srivastava *et al.* (1988) obtained the asymptotic variance of $\hat{\rho}_{ij,s}^{(ms)}$. It can be seen from Example 2 that the difference between the asymptotic variances of $\hat{\rho}_{ij,e}^{(ms)}$ and $\hat{\rho}_{ij,s}^{(ms)}$ depends upon only the values of $\omega_e - \omega_s$ for given $\rho_{ij}^{(ms)}$ and $\rho_{jj}^{(ss)}$. Srivastava and Keen (1988) showed that the asymptotic variance of $\hat{\rho}_{ij,s}^{(ms)}$ is smaller than that of $\hat{\rho}_{ij,e}^{(ms)}$.

In the univariate case where $p = q = 1$, $\hat{P}_{ms,p}$ in (3.8) and $\hat{P}_{ms,e}$ in (3.11) reduce to the pairwise and ensemble estimators, respectively. Konishi (1982) derived the asymptotic bias and variances of these estimators. The results have asymptotically the same form as those given in Examples 1 and 2.

Example 3. The estimator $\hat{P}_{ss,a}$ defined by (3.17) was constructed based on the unbiased estimators, in which the weights (3.16) can be rewritten as

$$\omega_a = 1 + \left(1 - \sum_\alpha k_\alpha^2 \Big/ \sum_\alpha k_\alpha\right) \Big/ \sum_\alpha (k_\alpha - 1) \quad \text{and}$$

$$v_a = -N \Big/ \sum_\alpha (k_\alpha - 1) + 1 \Big/ \sum_\alpha (k_\alpha - 1).$$

Hence the asymptotic distribution of $\sqrt{N}(\hat{P}_{ss,a} - P_{ss})$ is obtained by taking

$b_\alpha^{(s)} = k_\alpha$, $\omega_\alpha = 1$ and $\nu_\alpha = -N \Big/ \sum_\alpha (k_\alpha - 1)$ for $\alpha = 1, \ldots, N$ in Theorem 4.2.

Example 4. The estimator $\hat{P}_{ss,k}$ defined by (3.19) was obtained based on consistent estimators of Σ_s and Σ_{ss}. By an argument similar to that used for $\hat{P}_{ms,p}$ in Example 1, we can also apply Theorem 4.2 to derive the asymptotic distribution, in which the asymptotic variances and covariances are given by putting $N_s = U$, $b_\alpha^{(s)} = u_\alpha$ and taking ω_α and ν_α in (3.18). For the cases $k_\alpha = k$; $\alpha = 1, \ldots, N$ and $p = q = 1$, the pairwise intraclass correlation estimator reduces to the maximum likelihood estimator. Putting $k_\alpha = k$, $N_s = Nk(k-1)$, $b_\alpha^{(s)} = k(k-1)$, $\omega_\alpha = k-1$, $\nu_\alpha = -1$ for $\alpha = 1, \ldots, N$ and $i = j$ in (4.7), we have $2(1-\rho_{ii}^{(ss)})^2\{1+(k-1)\rho_{ii}^{(ss)}\}^2/\{k(k-1)\}$, which coincides with the asymptotic variance of the maximum likelihood estimator of the intraclass correlation coefficient (see Fisher (1958)).

4.2 *Interval estimation and hypothesis testing*

In order to assess the interrelationships among the different characteristics, we proposed the estimators $\hat{P}_{ms} = (\hat{\rho}_{ij}^{(ms)})$ and $\hat{P}_{ss} = (\hat{\rho}_{ij}^{(ss)})$ defined by (3.4) and (3.5), respectively. The distributional results in Theorems 4.1 and 4.2 can be used to construct approximate confidence intervals and test procedures for the interclass correlations $\rho_{ij}^{(ms)}$ and the intraclass correlations $\rho_{ij}^{(ss)}$.

It follows from Theorem 4.1 that the standardized quantity

$$\sqrt{N}(\hat{\rho}_{ij}^{(ms)} - \rho_{ij}^{(ms)})/\tau \tag{4.9}$$

is approximated by a standard normal distribution, where τ^2 is given by (4.1). To construct a confidence interval for $\rho_{ij}^{(ms)}$, we replace the unknown parameters $\rho_{ij}^{(ms)}$ and $\rho_{jj}^{(ss)}$ included in τ by their sample estimates. Then a confidence interval for $\rho_{ij}^{(ms)}$ with confidence coefficient $1-\alpha$ is approximately given by

$$\left(\hat{\rho}_{ij}^{(ms)} - \frac{1}{\sqrt{N}} z_{1-\alpha/2}\hat{\tau}, \hat{\rho}_{ij}^{(ms)} + \frac{1}{\sqrt{N}} z_{1-\alpha/2}\hat{\tau} \right),$$

where $\hat{\tau}$ is a consistent estimate of τ and $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$ percentile point of the standard normal distribution. As special cases of this result, we can construct approximate confidence intervals for $\rho_{ij}^{(ms)}$ based on each of the estimators given in Subsection 3.1.

The asymptotic result (4.9) can also be used for constructing tests of hypotheses. On the basis of multivariate observations from N families, we wish to test the hypothesis H_0: $\rho_{ij}^{(ms)} = \rho_0$ against H_1: $\rho_{ij}^{(ms)} \neq \rho_0$. Substituting a certain consistent estimate for $\rho_{jj}^{(ss)}$ and replacing $\rho_{ij}^{(ms)}$ by the specified value ρ_0 in τ, we obtain the consistent estimator, $\hat{\tau}_0$, of τ under the null hypothesis. Then a test statistic $\sqrt{N}(\hat{\rho}_{ij}^{(ms)} - \rho_0)/\hat{\tau}_0$ has approximately a standard normal distribution, so that H_0 would be rejected, at significance level α, if

$$|\sqrt{N}(\hat{\rho}_{ij}^{(ms)} - \rho_0)/\hat{\tau}_0| \geq z_{1-\alpha/2} .$$

A particular case of importance in the analysis of multivariate familial data is $\rho_0 = 0$. It is then required to test H_0: $\rho_{ij}^{(ms)} = 0$ against H_1: $\rho_{ij}^{(ms)} > 0$. An asymptotic test of size α is to reject H_0 if

$$\sqrt{N}\hat{\rho}_{ij}^{(ms)} \Bigg/ \left[N\left(\sum_{\alpha=1}^{N} b_\alpha^{(ms)} \right)^{-2} \sum_{\alpha=1}^{N} b_\alpha^{(ms)^2} \frac{1}{k_\alpha} \{1 + (k_\alpha - 1)\hat{\rho}_{jj}^{(ss)}\} \right]^{1/2} \geq z_{1-\alpha}$$

where $\hat{\rho}_{jj}^{(ss)}$ is a consistent estimate of $\rho_{jj}^{(ss)}$.

The asymptotic distribution of $\hat{\rho}_{ij}^{(ss)}$ given in Theorem 4.2 can be used to construct approximate confidence intervals and test procedures for the intraclass correlations. The intervals and test statistics are obtained in a similar manner to those for $\rho_{ij}^{(ms)}$.

Acknowledgements

The authors would like to thank Professor C. Radhakrishna Rao for helpful discussions and the referee for useful comments.

REFERENCES

Donner, A. (1986). A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model, *Internat. Statist. Rev.*, **54**, 67–82.

Donner, A. and Bull, S. (1984). A comparison of significance-testing procedures for parent-child correlations computed from family data, *Appl. Statist.*, **33**, 278–284.

Donner, A. and Eliasziw, M. (1988). Confidence interval construction for parent-offspring correlations, *Biometrics*, **44**, 727–737.

Donner, A. and Koval, J. J. (1980). The estimation of intraclass correlation in the analysis of family data, *Biometrics*, **36**, 19–25.

Fisher, R. A. (1958). *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.

Karlin, S., Cameron, E. C. and Williams, P. T. (1981). Sibling and parent-offspring correlation estimation with variable family size, *Proc. Nat. Acad. Sci. U.S.A.*, **78**, 2664–2668.

Konishi, S. (1982). Asymptotic properties of estimators of interclass correlation from familial data, *Ann. Inst. Statist. Math.*, **34**, 505–515.

Konishi, S. (1985). Testing hypotheses about interclass correlations from familial data, *Biometrics*, **41**, 167–176.

Rao, C. R. (1945). Familial correlations or the multivariate generalizations of the intraclass correlation, *Current Science*, **14**, 66–67.

Rao, C. R. (1953). Discriminant functions for genetic differentiation and selection, *Sankhyā*, **12**, 229–246.

Rosner, B. (1979). Maximum likelihood estimation of interclass correlations, *Biometrika*, **66**, 533–538.

Rosner, B., Donner, A. and Hennekens, C. H. (1977). Estimation of interclass correlation from familial data, *Appl. Statist.*, **26**, 179–187.

Srivastava, M. S. (1984). Estimation of interclass correlations in familial data, *Biometrika*, **71**, 177–185.

Srivastava, M. S. and Katapa, R. S. (1986). Comparison of estimators of interclass and intraclass correlations from familial data, *Canad. J. Statist.*, **14**, 29–42.

Srivastava, M. S. and Keen, K. J. (1988). Estimation of the interclass correlation coefficient, *Biometrika*, **75**, 731–739.

Srivastava, M. S., Keen, K. J. and Katapa, R. S. (1988). Estimation of interclass and intraclass correlations in multivariate familial data, *Biometrics*, **44**, 141–150.