

# Case Deletion Diagnostics for Intraclass Correlation Model

Myung Geun Kim<sup>1,a</sup>

<sup>a</sup>Department of Mathematics Education, Seowon University, Korea

---

## Abstract

The intraclass correlation model has a long history of applications in several fields of research. Case deletion diagnostic methods for the intraclass correlation model are proposed. Based on the likelihood equations, we derive a formula for a case deletion diagnostic method which enables us to investigate the influence of observations on the maximum likelihood estimates of the model parameters. Using the Taylor series expansion we develop an approximation to the likelihood distance. Numerical examples are provided for illustration.

**Keywords:** Case deletions, intraclass correlation, likelihood distance, Taylor series expansion.

---

## 1. Introduction

The intraclass correlation model also called equicorrelation model has a covariance matrix of equal variances and equal covariances among all the component variables. Therefore every pair of variables for the intraclass correlation model has the same correlation coefficient called the intraclass correlation coefficient, which was first introduced by Fisher (1921) in epidemiology to study familial resemblance between siblings with respect to some characteristics such as weight, height and blood pressure. Since then, the intraclass correlation coefficient has a lengthy history of applications in wide areas of research. In genetic linkage studies to quantify a proportion of variance explained by genetic factors, it is used to estimate the heritability of selected traits in animal and plant populations (Visscher, 1998; Visscher *et al.*, 2006). In psychology it measures the reliability of responses on one or more sets of judges or assessors (Ebel, 1951). In sensitivity analysis it is used to measure the effectiveness of an experimental treatment (Bradley and Schumann, 1957). In genomics it assesses the biological variation in DNA microarray analysis (Pellis *et al.*, 2003). A test of fit for the intraclass correlation model has been used for checking the validity of analysis-of-variance approach to repeated measure design (Renchner, 1995). Olkin and Pratt (1958) studied unbiased estimation of the intraclass correlation coefficient and some inferential results related with the intraclass correlation model can be found in Srivastava (1965), Krishnaiah and Pathak (1967), Bhargava and Srivastava (1973). More details about theory and methodology for inferences concerning the intraclass correlation coefficient can be found in Donner (1986), Donner and Koval (1980).

In this work, we will propose case deletion diagnostics for the intraclass correlation model. In Section 2, we will review maximum likelihood estimation of the model parameters, and provide some convenient form of the likelihood equations which makes it easy to derive case deletions. In Section 3, we will develop a case deletion diagnostic method which investigate the influence of observations on the maximum likelihood estimates of the model parameters. In Section 4, an approximation to the likelihood distance is developed using the Taylor series expansion. Numerical examples are given for illustration in Section 5. Concluding remarks are given in Section 6.

---

<sup>1</sup> Department of Mathematics Education, Seowon University, Cheongju 361-742, Korea. E-mail: [mgkim@seowon.ac.kr](mailto:mgkim@seowon.ac.kr)

## 2. Intraclass Correlation Model

We consider a random sample  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  from a  $p$ -variate normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$ , in which the covariance matrix has a pattern of equal variances and equal covariances. When we write as  $\mathbf{J}$  the  $p \times p$  matrix whose all elements are one, the intraclass correlation model has the covariance matrix defined by

$$\Sigma = \{(1 - \rho)\mathbf{I} + \rho\mathbf{J}\} \sigma^2,$$

where  $\rho$  is a real number and  $\sigma > 0$ . The determinant of  $\Sigma$  is given by

$$|\Sigma| = \{1 + (p - 1)\rho\}(1 - \rho)^{p-1} \sigma^{2p}.$$

If  $\rho \neq 1$  or  $\rho \neq -(p - 1)^{-1}$ , then the inverse of  $\Sigma$  exists and it is given by

$$\Sigma^{-1} = [\mathbf{I} - \rho\{1 + (p - 1)\rho\}^{-1}\mathbf{J}] \{\sigma^2(1 - \rho)\}^{-1}.$$

More details about the equicorrelation matrix can be found in Rencher (1995) and Mardia *et al.* (1979).

Based on the full sample of size, the sample mean vector and the sample covariance matrix are given by

$$\begin{aligned} \bar{\mathbf{x}} &= \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i, \\ \mathbf{S} = (s_{ij}) &= \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \end{aligned}$$

respectively. The maximum likelihood estimators of the model parameters are given by

$$\begin{aligned} \hat{\mu} &= \bar{\mathbf{x}}, \\ \hat{\sigma}^2 &= \frac{1}{p} \text{tr}(\mathbf{S}), \\ \hat{\rho} \hat{\sigma}^2 &= \frac{2}{p(p-1)} \sum_{i < j} s_{ij} = \frac{1}{p(p-1)} \{\mathbf{1}_p^T \mathbf{S} \mathbf{1}_p - \text{tr}(\mathbf{S})\}, \end{aligned}$$

where  $\mathbf{1}_p$  is the  $p \times 1$  vector whose all elements are one. Some materials about maximum likelihood estimation for the intraclass correlation model can be found in Morrison (1990) and Rencher (1995). The second identity for  $\hat{\rho} \hat{\sigma}^2$  enables us to easily derive a formula for case deletions, which will be treated later.

## 3. Case Deletions

In what follows, the quantity with subscript  $(r)$  indicates that the corresponding quantity is computed without the  $r$ th observation. The influence of  $\mathbf{x}_r$  on the estimate of each model parameter can be investigated using  $\hat{\mu} - \hat{\mu}_{(r)}$ ,  $\hat{\sigma}^2 - \hat{\sigma}_{(r)}^2$  and  $\hat{\rho} - \hat{\rho}_{(r)}$ , which will be derived in this section.

Based on the reduced sample of size  $n - 1$  without the  $r^{th}$  observation  $\mathbf{x}_r$ , the sample mean vector and the sample covariance matrix are given by

$$\bar{\mathbf{x}}_{(r)} = \frac{1}{n-1} \sum_{i=1, i \neq r}^n \mathbf{x}_i = \frac{n}{n-1} \bar{\mathbf{x}} - \frac{1}{n-1} \mathbf{x}_r,$$

$$\mathbf{S}_{(r)} = \frac{1}{n-1} \sum_{i=1, i \neq r}^n (\mathbf{x}_i - \bar{\mathbf{x}}_{(r)}) (\mathbf{x}_i - \bar{\mathbf{x}}_{(r)})^T = \frac{n}{n-1} \mathbf{S} - \frac{n}{(n-1)^2} (\mathbf{x}_r - \bar{\mathbf{x}}) (\mathbf{x}_r - \bar{\mathbf{x}})^T,$$

respectively. For ease of presentations we let

$$c_r = (\mathbf{x}_r - \bar{\mathbf{x}})^T (\mathbf{x}_r - \bar{\mathbf{x}}),$$

$$d_r = \left\{ \mathbf{1}_p^T (\mathbf{x}_r - \bar{\mathbf{x}}) \right\}^2.$$

From the likelihood equations based on the full sample given in the previous section we can surmise the form of the likelihood equations based on the reduced sample without the  $r$ th observation which turn out to be

$$\hat{\mu}_{(r)} = \bar{\mathbf{x}}_{(r)},$$

$$\hat{\sigma}_{(r)}^2 = \frac{1}{p} \text{tr}(\mathbf{S}_{(r)}),$$

$$\hat{\rho}_{(r)} \hat{\sigma}_{(r)}^2 = \frac{1}{p(p-1)} \left\{ \mathbf{1}_p^T \mathbf{S}_{(r)} \mathbf{1}_p - \text{tr}(\mathbf{S}_{(r)}) \right\}$$

from which we easily get

$$\hat{\mu} - \hat{\mu}_{(r)} = \frac{1}{n-1} (\mathbf{x}_r - \bar{\mathbf{x}}),$$

$$\hat{\sigma}^2 - \hat{\sigma}_{(r)}^2 = -\frac{1}{n-1} \hat{\sigma}^2 + \frac{n}{p(n-1)^2} c_r,$$

$$(\hat{\rho} - \hat{\rho}_{(r)}) \hat{\sigma}_{(r)}^2 = -\frac{n}{p(n-1)^2} \left( \hat{\rho} + \frac{1}{p-1} \right) c_r + \frac{1}{p(p-1)} \frac{n}{(n-1)^2} d_r.$$

From the second equation in the above, we can find  $\hat{\sigma}_{(r)}^2$  with which the third equation above yield a formula for  $\hat{\rho} - \hat{\rho}_{(r)}$ . Giraudeau *et al.* (1996) provided a formula for case deletion of the intraclass correlation coefficient only, from which the form of our derivation based on the likelihood equations themselves is different.

#### 4. Measures based on the Likelihood

Let  $\theta = (\mu^T, \sigma^2, \rho)^T$  and  $l(\theta)$  be the log-likelihood function of  $\theta$  based on the full sample. The influence of the  $r$ th observation on the likelihood can be measured using the likelihood distance (Cook and Weisberg, 1982; Chatterjee and Hadi, 1988) defined by

$$LD_r = 2 \{l(\theta) - l(\theta_{(r)})\}.$$

The Taylor series expansion of  $l(\theta_{(r)})$  about  $\hat{\theta}$  approximates  $LD_r$  by

$$LD_r \approx \left( \hat{\theta} - \hat{\theta}_{(r)} \right)^T \ddot{l}(\hat{\theta}) \left( \hat{\theta} - \hat{\theta}_{(r)} \right), \quad (4.1)$$

where  $\ddot{l}(\hat{\theta}) = \partial^2 l(\theta) / \partial \theta \partial \theta^T$  evaluated at  $\theta = \hat{\theta}$ . More details about approximations by Taylor series expansion can be found in Seber and Wild (1989).

The log-likelihood  $l(\theta)$  can be written, ignoring unimportant constant, as

$$l(\theta) = -\frac{n}{2} \left[ \log\{1 + (p-1)\rho\} + (p-1) \log(1-\rho) + p \log(\sigma^2) \right] \\ - \frac{1}{2} \frac{1}{\sigma^2(1-\rho)} \sum_{i=1}^n (\mathbf{x}_i - \mu)^T (\mathbf{x}_i - \mu) - \frac{1}{2} \frac{\rho}{\{1 + (p-1)\rho\}\sigma^2(1-\rho)} \sum_{i=1}^n \{\mathbf{1}_p^T (\mathbf{x}_i - \mu)\}^2.$$

In order to get an approximation to  $LD_r$  given in (4.1), we need first to compute  $\dot{l}(\hat{\theta}) = \partial l(\theta) / \partial \theta$  whose components are easily obtained as

$$\frac{\partial l(\theta)}{\partial \mu} = n \left[ I - \rho \{1 + (p-1)\rho\}^{-1} J \right] \{\sigma^2(1-\rho)\}^{-1} (\bar{\mathbf{x}} - \mu), \\ \frac{\partial l(\theta)}{\partial (\sigma^2)} = -\frac{np}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{1}{\sigma^4(1-\rho)} \sum_{i=1}^n (\mathbf{x}_i - \mu)^T (\mathbf{x}_i - \mu) \\ - \frac{1}{2} \frac{\rho}{\{1 + (p-1)\rho\}\sigma^4(1-\rho)} \sum_{i=1}^n \{\mathbf{1}_p^T (\mathbf{x}_i - \mu)\}^2, \\ \frac{\partial l(\theta)}{\partial (\rho)} = \frac{n}{2} \frac{p(p-1)\rho}{\{1 + (p-1)\rho\}(1-\rho)} - \frac{1}{2} \frac{1}{\sigma^2(1-\rho)^2} \sum_{i=1}^n (\mathbf{x}_i - \mu)^T (\mathbf{x}_i - \mu) \\ + \frac{1}{2} \frac{1 + (p-1)\rho^2}{\{1 + (p-1)\rho\}^2 \sigma^2(1-\rho)^2} \sum_{i=1}^n \{\mathbf{1}_p^T (\mathbf{x}_i - \mu)\}^2.$$

A little computation shows that the following identities hold

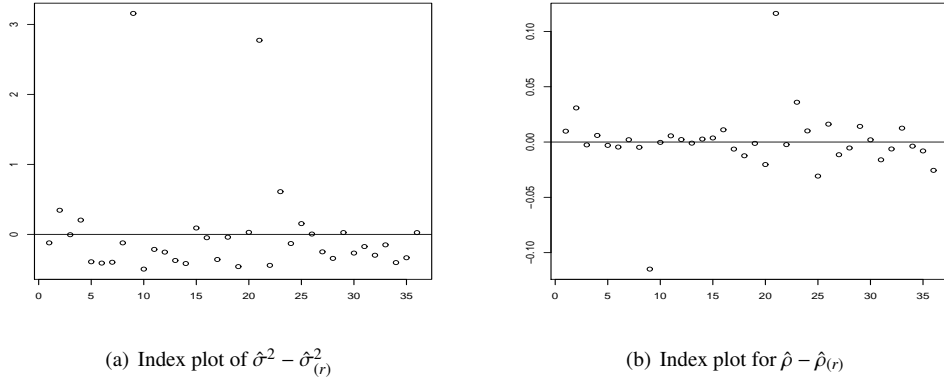
$$\sum_{i=1}^n d_r = n \text{tr}(\mathbf{S}) + 2n \sum_{i < j} s_{ij} = n \mathbf{1}_p^T \mathbf{S} \mathbf{1}_p = n\{1 + (p-1)\hat{\rho}\} \text{tr}(\mathbf{S})$$

with whose aids we easily get the components of  $\ddot{l}(\hat{\theta})$  as follows

$$\frac{\partial^2 l(\theta)}{\partial \mu \partial \mu^T} = -n \left[ I - \hat{\rho} \{1 + (p-1)\hat{\rho}\}^{-1} J \right] \{\hat{\sigma}^2(1-\hat{\rho})\}^{-1}, \\ \frac{\partial^2 l(\theta)}{\partial \mu \partial (\sigma^2)} = \frac{\partial^2 l(\theta)}{\partial \mu \partial \rho} = 0, \\ \frac{\partial^2 l(\theta)}{\partial (\sigma^2)^2} = -\frac{np}{2} \frac{1}{\hat{\sigma}^4}, \\ \frac{\partial^2 l(\theta)}{\partial \rho \partial (\sigma^2)} = \frac{np(p-1)}{2} \frac{\hat{\rho}}{\{1 + (p-1)\hat{\rho}\}(1-\hat{\rho})\hat{\sigma}^2}, \\ \frac{\partial^2 l(\theta)}{\partial \rho^2} = -\frac{np(p-1)}{2} \frac{1 + (p-1)\hat{\rho}^2}{(1-\hat{\rho})^2 \{1 + (p-1)\hat{\rho}\}^2}.$$

Together with the formula for case deletions obtained in Section 3, we are in a position to get an approximation to  $LD_r$  given in (4.1) given by

$$LD_r \approx A_1 + A_2 + A_3 + A_4 + A_5,$$

Figure 1: Case deletions for  $\hat{\sigma}^2$  and  $\hat{\rho}$ : cost data

where

$$\begin{aligned}
 A_1 &= \frac{np}{2(n-1)^2} + \frac{n}{(n-1)^3} \frac{n\hat{\rho} - 1}{1 - \hat{\rho}} \frac{c_r}{\hat{\sigma}^2} - \frac{n}{(n-1)^2} \frac{\hat{\rho}}{(1 - \hat{\rho})\{1 + (p-1)\hat{\rho}\}} \frac{d_r}{\hat{\sigma}^2}, \\
 A_2 &= \frac{n^3}{2p(n-1)^4} \frac{c_r^2}{\hat{\sigma}^4} - \frac{n^2}{(n-1)^3} \frac{\hat{\rho}}{1 - \hat{\rho}} \frac{c_r}{\hat{\sigma}_{(r)}^2} + \frac{n^3}{p(n-1)^4} \frac{\hat{\rho}}{1 - \hat{\rho}} \frac{1}{\hat{\sigma}^2} \frac{c_r^2}{\hat{\sigma}_{(r)}^2}, \\
 A_3 &= \frac{n^2}{(n-1)^3} \frac{\hat{\rho}}{1 - \hat{\rho}} \frac{1}{1 + (p-1)\hat{\rho}} \frac{d_r}{\hat{\sigma}_{(r)}^2} - \frac{n^3}{p(n-1)^4} \frac{\hat{\rho}}{1 - \hat{\rho}} \frac{1}{1 + (p-1)\hat{\rho}} \frac{1}{\hat{\sigma}^2} \frac{c_r d_r}{\hat{\sigma}_{(r)}^2}, \\
 A_4 &= \frac{n^3}{2p(p-1)(n-1)^4} \frac{1 + (p-1)\hat{\rho}^2}{(1 - \hat{\rho})^2} \frac{c_r^2}{\hat{\sigma}_{(r)}^4} - \frac{n^3}{p(p-1)(n-1)^4} \frac{1 + (p-1)\hat{\rho}^2}{(1 - \hat{\rho})^2\{1 + (p-1)\hat{\rho}\}} \frac{c_r d_r}{\hat{\sigma}_{(r)}^4}, \\
 A_5 &= \frac{n^3}{2p(p-1)(n-1)^4} \frac{1 + (p-1)\hat{\rho}^2}{(1 - \hat{\rho})^2\{1 + (p-1)\hat{\rho}\}^2} \frac{d_r^2}{\hat{\sigma}_{(r)}^4}.
 \end{aligned}$$

Note that in computing an approximation to  $LD_r$ ,  $\hat{\sigma}_{(r)}^2$  can be rewritten as

$$\hat{\sigma}_{(r)}^2 = \frac{n}{p(n-1)^2} \{p(n-1)\hat{\sigma}^2 - c_r\}.$$

## 5. Numerical Examples

For illustration the intraclass correlation model is fitted to two data sets. First, we consider the cost data with three variables (Johnson and Wichern, 1992, p.276) for illustrating the diagnostic method developed in the previous sections. The maximum likelihood estimates for the model parameters based on the full sample are  $\hat{\mu} = (12.22, 8.11, 9.59)^T$ ,  $\hat{\sigma}^2 = 17.67$  and  $\hat{\rho} = 0.37$ . We will consider case deletions for  $\hat{\sigma}^2$ ,  $\hat{\rho}$  and likelihood distances.

An index plot for the values  $\hat{\sigma}^2 - \hat{\sigma}_{(r)}^2$  is displayed in Figure 1(a), which shows that observations 9 and 21 are influential in estimating  $\sigma^2$ . Deletion of each of observations 9 and 21 moves  $\hat{\sigma}^2 = 17.67$  down to 14.51 and 14.90, respectively, and in this respect the role of deleting observations 9 is similar to that of deleting observation 21. An index plot for the values  $\hat{\rho} - \hat{\rho}_{(r)}$  is included in Figure 1(b), which shows that observations 9 and 21 are influential also in estimating  $\rho$ . The role of observations 9 and 21 for the case of  $\hat{\rho}$  is different from that for the case of  $\hat{\sigma}^2$  because deletion of observation 9 increases  $\hat{\rho} = 0.37$  to 0.48 while deletion of observation 21 decreases  $\hat{\rho}$  to 0.25.

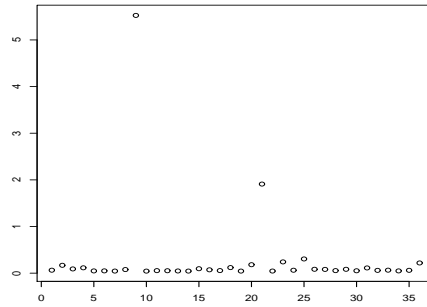
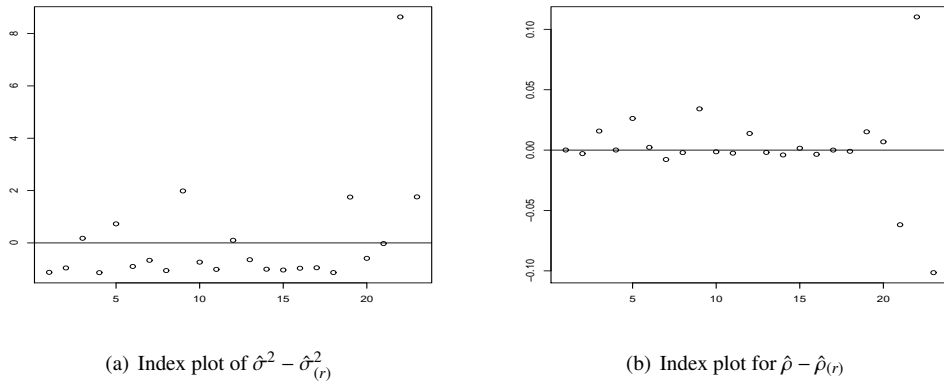
Figure 2: Index plot for  $LD_r$ : cost dataFigure 3: Case deletions for  $\hat{\sigma}^2$  and  $\hat{\rho}$ : simulated data

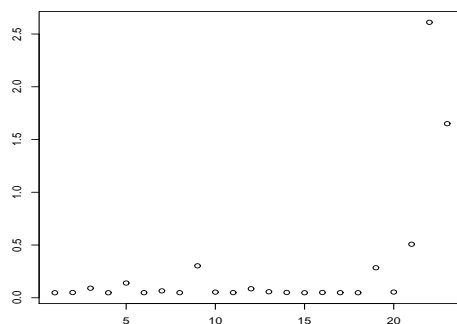
Figure 2 shows an index plot for the values of  $LD_r$ , from which we can see that deletion of observations 9 has the largest influence on the likelihood and observation 21 is the next one. The influence of other observations except for these two observations on the likelihood may be neglected. Observations 9 and 21 stand out in all of three displays.

Next, the intraclass correlation model is fitted to the simulated data with 23 observations on two variables (Kim, 2000) in which observations 21, 22, 23 have large influence on the variance and the correlation. The maximum likelihood estimates for the model parameters based on the full sample are  $\hat{\mu} = (1.00, 0.19)^T$ ,  $\hat{\sigma}^2 = 25.28$  and  $\hat{\rho} = 0.61$ .

An index plot for the values  $\hat{\sigma}^2 - \hat{\sigma}_{(r)}^2$  in Figure 3(a) shows that observation 22 is remarkably influential in estimating  $\sigma^2$  compared with the others. Deletion of observations 22 moves  $\hat{\sigma}^2 = 25.28$  down to 16.65. An index plot for the values  $\hat{\rho} - \hat{\rho}_{(r)}$  in Figure 3(b) shows that observations 21, 22, 23 are influential in estimating  $\rho$ . Figure 4 shows an index plot for the values of  $LD_r$ , from which we can see that deletion of observations 22 has the largest influence on the likelihood and observation 23, 21 are the next.

## 6. Concluding Remarks

A direct derivation of Cook's distance for  $\hat{\theta} - \hat{\theta}_{(r)}$  is not feasible because the covariance matrix of  $\hat{\theta}$  is not at present available. The covariance matrix of the asymptotic distribution of the maximum likelihood estimator  $\hat{\theta}$  is just  $-\{E(\partial^2 l(\theta)/\partial \theta \partial \theta^T)\}^{-1}$ . Hence if  $\partial^2 l(\theta)/\partial \theta \partial \theta^T$  is a good approximation to

Figure 4: Index plot for  $LD_r$ : simulated data

$E(\partial^2 l(\theta)/\partial\theta\partial\theta^T)$  at  $\theta = \hat{\theta}$ , then  $LD_r$  in (4.1) can be used as an approximation to Cook's distance for  $\hat{\theta} - \hat{\theta}_{(r)}$ .

Though this work centers on a single case deletion, multiple case deletions can be performed, which may be further work. Case deletion diagnostic suffers from so called masking and swamping effects, which may be avoided by a robust method that is also a further work.

## References

- Bhargava, R. P. and Srivastava, M. S. (1973). On Tukey's confidence intervals for the contrasts in the means of the intraclass correlation model, *Journal of the Royal Statistical Society, Series B*, **35**, 147–152.
- Bradley, R. A. and Schumann, D. E. W. (1957). The comparison of the sensitivities of similar experiments: Applications, *Biometrics*, **13**, 496–510.
- Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression*, Wiley, New York.
- Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression*, Chapman and Hall, New York.
- Donner, A. (1986). A review of inference procedures for the intraclass correlation coefficient on the one-way random effects model, *International Statistical Review*, **54**, 67–82.
- Donner, A. and Koval, J. J. (1980). The large sample variance of an intraclass correlation, *Biometrika*, **67**, 719–722.
- Ebel, R. L. (1951). Estimation of the reliability of ratings, *Psychometrika*, **16**, 407–424.
- Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample, *Metron*, **1**, 1–32.
- Giraudeau, B., Mallet, A. and Chastang, C. (1996). Case influence on the intraclass correlation coefficient estimate, *Biometrics*, **52**, 1492–1497.
- Johnson, A. J. and Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis*, 3rd ed., Prentice-Hall.
- Kim, M. G. (2000). Multivariate outliers and decompositions of Mahalanobis distance, *Communications in Statistics: Theory and Methods*, **29**, 1511–1526.
- Krishnaiah, P. R. and Pathak, P. K. (1967). Tests for the equality of covariance matrices under the intraclass correlation model, *The Annals of Mathematical Statistics*, **38**, 1286–1288.
- Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*, Academic Press.
- Morrison, D. F. (1990). *Multivariate Statistical Methods*, 3rd ed., McGraw-Hill.

- Olkin, I. and Pratt, J. W. (1958). Unbiased estimation of certain correlation coefficients, *The Annals of Mathematical Statistics*, **29**, 201–211.
- Pellis, L., Hal, N. L., von Burema, J. and Keijer, J. (2003). The intraclass correlation coefficient applied for evaluation of data correction, labeling methods, and rectal biopsy sampling in DNA microarray experiments, *Physiological Genomics*, **16**, 99–106.
- Rencher, A. C. (1995). *Methods of Multivariate Analysis*, Wiley, New York.
- Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*, Wiley, New York.
- Srivastava, M. S. (1965). Some tests for the intraclass correlation model, *The Annals of Mathematical Statistics*, **36**, 1802–1806.
- Visscher, P. M. (1998). On the sampling variance of intraclass correlations and genetic correlation, *Genetics*, **149**, 1605–1614.
- Visscher, P. M., Medland, S. E., Ferreira, M. A. R., Morley, K. I., Zhu, G., Cornes, B. K., Montgomery, G. W. and Martin, N. G. (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings, *PLoS Genetics*, **2**, e41.

Received March 12, 2014; Revised April 2, 2014; Accepted April 2, 2014