On testing equality of intraclass correlations under unequal family sizes

Dayanand N. Naik*, Amal Helu

Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA 23529, USA

Received 16 May 2006; received in revised form 27 February 2007; accepted 28 February 2007
Available online 12 March 2007

Abstract
The three likelihood-based tests, namely, likelihood ratio test, Rao score test, and Wald test and two more asymptotic tests which use Srivastava’s estimator of intraclass correlation coefficient are considered to test the null hypothesis of equality of intraclass correlation coefficients when the families have unequal number of children. Methods are illustrated on Galton’s data set. Using simulation experiment we compute the sizes and powers of these tests and compare. It is found that our proposed test using Srivastava’s estimator and the score test perform the best among all tests.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Intraclass correlation; Likelihood ratio test; Rao score test; Srivastava estimator; Wald test

1. Introduction
To measure the degree of resemblance between family members with respect to a specified characteristic, sib–sib and parent–sib correlations are used. Testing equality of sib–sib or intraclass correlation coefficients from different populations is an important problem in the study of familial correlations. This problem for two populations was considered by Donner and Bull (1983) when family sizes within a population and between populations were the same and by Khatri et al. (1989) when family sizes in the two populations were different. These authors derived and studied the performance of the likelihood ratio test (LRT).

For the problem of testing equality of several correlations, Konishi and Gupta (1989) have suggested a modified LRT and a test based on Fisher’s z-transformation, Paul and Barnwal (1990) suggested a C(z) test, and Haung and Sinha (1993) derived the optimum invariant test, assuming the family sizes within population are the same, but different for different populations.

Young and Bhandary (1998) and Bhandary and Alam (2000), respectively, considered the problems of testing the equality of two and three correlation coefficients when the family sizes are unequal. They used Srivastava’s (1984) estimator of intraclass correlation and proposed the approximate LRT and compared its performance with two other
asymptotic tests based on normal distribution. They also made the assumption that the variances for different populations are the same.

Recently, Hanley (2004) brought Galton’s familial data on human stature into the light. This is a very interesting familial data set. One of the problems of interest may be to compare the intraclass correlation coefficient between boys in a family with that between girls in another family. Then the families from which only boys are considered forms one population for the study and the families from which only girls are considered forms another population. Under this situation there is no reason to believe that the two variances of the two populations are equal.

In this article, we consider the problem of testing equality of \( k \) intraclass correlations when the family sizes are unequal and the variances different for different populations. For testing, we propose to use maximum likelihood-based three asymptotic tests, namely, the LRT, Wald test, and Rao score test. In the spirit of Young and Bhandary (1998) we also propose two more asymptotic tests. A comparison of all these tests is made using simulation experiments. In the next section we provide the maximum likelihood-based tests and in Section 3 Srivastava’s estimators under null and non-null cases are given. In Section 4, two asymptotic tests will be provided based on Srivastava’s estimate and the next section we provide the maximum likelihood-based tests and in Section 6, we will describe the simulation experiment and provide the results and conclusions.

2. Likelihood-based asymptotic tests

Suppose there are \( g \) independent populations (or groups) and data on children of the families randomly selected from the populations are available. Suppose in the \( i \)th population there are \( n_i \) families and the number of children in families is allowed to be different. We denote the number of children in the \( j \)th family from the \( i \)th population by \( m_{ij} \).

Suppose \( x_{ijk}, k = 1, \ldots, m_{ij}; j = 1, \ldots, n_i; i = 1, \ldots, g \) are the observations on the \( k \)th child of the \( j \)th family belonging to the \( i \)th population. We assume for the \( i \)th population that \( E(x_{ijk}) = \mu_i \), \( \text{var}(x_{ijk}) = \sigma_i^2 \), and the sib–sib or intraclass correlation, \( \text{corr}(x_{ijk}, x_{ijk'}) = \rho_i \) for \( k \neq k' \). For every \( i \), we have \( -\infty < \mu_i < \infty \), \( \sigma_i^2 > 0 \), and \( -1/\max_{1 \leq j \leq n_i} m_{ij} - 1 < \rho_i < 1 \). Let the vector of observations on the \( j \)th family from \( i \)th group be \( x_{ij} = (x_{ij1}, \ldots, x_{ijm_{ij}})' \). Then \( E(x_{ij}) = \mu_{ij} = \mu_i 1_{m_{ij}} \), where \( 1_{m} \) (in general) is an \( m \times 1 \) vector of all ones, and the variance–covariance matrix of \( x_{ij} \) is

\[
var(x_{ij}) = \Sigma_{ij} = \sigma_i^2 [(1 - \rho_i) 1_{m_{ij}} + \rho_i J_{m_{ij}}]
\]

\[
= \sigma_i^2 V_{ij}(\rho_i) = \sigma_i^2 \begin{pmatrix}
1 & \rho_i & \cdots & \rho_i \\
\rho_i & 1 & \cdots & \rho_i \\
\vdots & \vdots & \ddots & \vdots \\
\rho_i & \rho_i & \cdots & 1
\end{pmatrix},
\]

where \( 1_{m} \) is an identity matrix of order \( m \) and \( J_{m_{ij}} \) is the \( m \times m \) matrix of all ones.

Note that the determinant and inverse of \( \Sigma_{ij} \), respectively, are

\[
|\Sigma_{ij}| = (\sigma_i^2)^{m_{ij}} [(1 - \rho_i)^{m_{ij} - 1} (1 + (m_{ij} - 1) \rho_i)]
\]

and

\[
\Sigma_{ij}^{-1} = \frac{1}{\sigma_i^2 (1 - \rho_i)} \left[ 1_{m_{ij}} - \rho_i \frac{1}{1 + (m_{ij} - 1) \rho_i} J_{m_{ij}} \right].
\]

Let us assume that \( x_{ij} \sim N_{m_{ij}}(\mu_{ij}, \Sigma_{ij}) \), \( j = 1, \ldots, n_i; \ i = 1, \ldots, g \), that is, the vector of observations \( x_{ij} \) is distributed as an \( m_{ij} \)-variate normal with mean vector \( \mu_{ij} \) and variance–covariance matrix \( \Sigma_{ij} \).

Suppose \( \theta = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho_1, \sigma_2^2, \rho_2, \ldots, \mu_g, \sigma_g^2, \rho_g)' \). Then the likelihood function can be written as

\[
L(\theta) = \prod_{i=1}^{g} \prod_{j=1}^{n_i} \frac{1}{(2\pi)^{m_{ij}/2}|\Sigma_{ij}|^{1/2}} \exp \left\{ -\frac{1}{2} (x_{ij} - \mu_{ij})' \Sigma_{ij}^{-1} (x_{ij} - \mu_{ij}) \right\}.
\]
More explicitly, the log-likelihood function can be expressed as

\[
\log L(\theta) = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \left\{ -\frac{m_{ij}}{2} \log(2\pi \sigma_i^2) - \frac{1}{2} (m_{ij} - 1) \log(1 - \rho_i) + \log(1 + (m_{ij} - 1)\rho_i) \right\} \\
- \frac{1}{2\sigma_i^2(1 - \rho_i)} (x_{ij} - \mu_i \mathbf{1}_{m_{ij}})' \left[ \mathbf{I}_{m_{ij}} - \frac{\rho_i}{1 + (m_{ij} - 1)\rho_i} \mathbf{J}_{m_{ij}} \right] (x_{ij} - \mu_i \mathbf{1}_{m_{ij}}).
\]

Let \( \hat{\theta} \) be the maximum likelihood estimator (MLE) of \( \theta \) which is obtained by maximizing \( L(\theta) \) and \( \hat{\theta}_0 \) be the MLE obtained by maximizing \( L(\theta) \) under the null hypothesis \( H_0 : \rho_1 = \rho_2 = \cdots = \rho_g \). Note that \( \theta_0 = (\mu_1, \sigma_1^2, \rho, \mu_2, \sigma_2^2, \rho, \ldots, \mu_g, \sigma_g^2, \rho)' \), which is same as \( \theta \), but under the null hypothesis.

2.1. The likelihood ratio test

The LRT statistic is

\[
LRT = -2 \log(A) = 2 \log L(\hat{\theta}) - 2 \log L(\hat{\theta}_0).
\] (1)

We reject the null hypothesis for the large values and the asymptotic distribution of the test statistic is \( \chi^2 \) with \( g - 1 \) degrees of freedom.

2.2. Rao’s score test

Let \( S(\theta) = \hat{\theta} \log L(\theta)/\hat{\theta} \) be \( 3g \times 1 \) vector of the score function and \( S(\theta) = E[(\hat{\theta} \log L(\theta)/\hat{\theta})(\hat{\theta} \log L(\theta)/\hat{\theta})'] \) be the \( 3g \times 3g \) Fisher information matrix.

Then the score test statistic is

\[
Score = S(\hat{\theta}_0)'S(\hat{\theta}_0)^{-1}S(\hat{\theta}_0).
\] (2)

The asymptotic distribution of this is \( \chi^2 \) with \( g - 1 \) degrees of freedom.

Note that \( S(\theta) = (S_1(\theta), \ldots, S_g(\theta))' \), where

\[
S_j(\theta) = (\hat{\theta} \log L(\theta)/\hat{\theta} \mu_j, \hat{\theta} \log L(\theta)/\hat{\theta} \sigma_j^2, \hat{\theta} \log L(\theta)/\hat{\theta} \rho_j)' \]

and

\[
\frac{\hat{\theta} \log L(\theta)}{\hat{\theta} \mu_j} = \frac{1}{\sigma_j^2(1 - \rho_j)} \sum_{j=1}^{n_j} \mathbf{1}_{m_{ij}}' \left[ \mathbf{I}_{m_{ij}} - \frac{\rho_i}{1 + (m_{ij} - 1)\rho_i} \mathbf{J}_{m_{ij}} \right] \mathbf{e}_{ij},
\]

\[
\frac{\hat{\theta} \log L(\theta)}{\hat{\theta} \sigma_j^2} = -\frac{1}{2\sigma_j^2} \sum_{j=1}^{n_j} m_{ij} + \sum_{j=1}^{n_j} \frac{1}{2\sigma_j^4(1 - \rho_j)} \mathbf{e}_{ij}' \left[ \mathbf{I}_{m_{ij}} - \frac{\rho_i}{1 + (m_{ij} - 1)\rho_i} \mathbf{J}_{m_{ij}} \right] \mathbf{e}_{ij},
\]

\[
\frac{\hat{\theta} \log L(\theta)}{\hat{\theta} \rho_j} = \frac{\rho_i}{2(1 - \rho_i)} \sum_{j=1}^{n_j} m_{ij}(m_{ij} - 1)
- \frac{1}{2\sigma_j^2(1 - \rho_j)^2} \sum_{j=1}^{n_j} \mathbf{e}_{ij}' \left[ \mathbf{I}_{m_{ij}} - \frac{1 + (m_{ij} - 1)\rho_i^2}{(1 + (m_{ij} - 1)\rho_i)^2} \mathbf{J}_{m_{ij}} \right] \mathbf{e}_{ij},
\]

where \( \mathbf{e}_{ij} = x_{ij} - \mu_i \mathbf{1}_{m_{ij}} \).

Next, we note that Fisher information matrix \( S(\theta) \) is a block diagonal matrix containing \( g \) blocks of \( 3 \times 3 \) matrices. Then \( S(\theta)^{-1} \), the inverse of Fisher information matrix will also be block diagonal. In the following we provide the
expressions for the \( i \)th block, \( \mathcal{S}_i \), of Fisher information matrix which can be used in practice and inverse of which can be computed.

\[
\mathcal{S}_i = \begin{pmatrix}
\sum_{j=1}^{n_i} \frac{m_{ij}}{\sigma_i^2(1+(m_{ij}-1)\rho_i)} & 0 & 0 \\
0 & \sum_{j=1}^{n_i} \frac{m_{ij}}{2\sigma_i^2} & 0 \\
0 & 0 & \sum_{j=1}^{n_i} \frac{\rho_j m_{ij}(m_{ij}-1)}{2(1-\rho_j)(1+(m_{ij}-1)\rho_j)}
\end{pmatrix}.
\]

2.3. Wald’s test

Let \( \mathbf{\rho} = (\rho_1, \rho_2, \ldots, \rho_g)' \) and \( \hat{\mathbf{\rho}} = (\hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_g)' \) be the MLE of \( \mathbf{\rho} \). Then the null hypothesis \( H_0 : \rho_1 = \rho_2 = \cdots = \rho_g \) can be expressed as \( \mathbf{C}\mathbf{\rho} = 0 \), where \( \mathbf{C} \) is a \((g-1) \times g\) contrast matrix, for example,

\[
\mathbf{C} = \begin{pmatrix}
1 & -1 & 0 & \cdots & 0 & 0 \\
0 & 1 & -1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & -1
\end{pmatrix}.
\] (3)

Let \( \mathbf{V}_\mathbf{\rho} \) be the \( g \times g \) diagonal matrix constructed from \( g \) blocks of \( \mathcal{S}(\hat{\mathbf{\theta}})^{-1} \) by sequentially (starting with the first block) taking only the third diagonal element from each block. Then the Wald’s test statistic is

\[
Wald = (\mathbf{C}\hat{\mathbf{\rho}})'[\mathbf{C}\mathbf{V}_\mathbf{\rho}\mathbf{C}']^{-1}(\mathbf{C}\hat{\mathbf{\rho}}).
\] (4)

The asymptotic distribution of this statistic is also \( \chi^2 \) with \( g-1 \) degrees of freedom.

In our simulation experiment we have used optimization routines in SAS’ IML procedure. The maximum likelihood estimates \( \hat{\mathbf{\theta}} \) and \( \hat{\mathbf{\theta}}_0 \) are easy to obtain by directly maximizing the likelihood function \( L(\mathbf{\theta}) \) or \( \log L(\mathbf{\theta}) \) under no restrictions and under the restrictions of the null hypothesis, respectively.

In the next section we introduce Srivastava’s non-iterative estimators of the unknown parameters and propose modifications to the above three tests by substituting Srivastava’s estimators in place of the maximum likelihood estimates in the expressions for computing the test statistics.

3. Srivastava’s estimators

Noting that the computation of the maximum likelihood estimates, Srivastava (1984) suggested alternative estimators to maximum likelihood estimators. Suppose we transform the vector of observations, \( \mathbf{x}_{ij} = (x_{ij1}, \ldots, x_{ijm_{ij}})' \) for the \( j \)th family, \( j = 1, \ldots, n_i \) and \( i \)th population, \( i = 1, \ldots, g \) to

\[
\mathbf{y}_{ij} = (y_{ij1}, y_{ij2}, \ldots, y_{ijm_{ij}})' = \mathbf{\Gamma}_{ij} \mathbf{x}_{ij},
\]

where \( \mathbf{\Gamma}_{ij} \) is an \( m_{ij} \times m_{ij} \) orthogonal matrix, say Helmert matrix (with a slight modification to the first row) given by

\[
\mathbf{\Gamma}_{ij} = \begin{pmatrix}
\frac{1}{m_{ij}} & \frac{1}{m_{ij}} & \frac{1}{m_{ij}} & \cdots & \frac{1}{m_{ij}} \\
1 & -1 & 0 & \cdots & 0 \\
\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & -2 & \cdots & 0 \\
\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{\sqrt{m_{ij}(m_{ij}-1)}} & \frac{1}{\sqrt{m_{ij}(m_{ij}-1)}} & \frac{1}{\sqrt{m_{ij}(m_{ij}-1)}} & \cdots & \frac{-1}{\sqrt{m_{ij}(m_{ij}-1)}}
\end{pmatrix}.
\]
It is clear that in this transformation, $y_{ij1}$ is the average of the sib scores and $y_{ijk}, \ k=2, \ldots, m_{ij}$ are the transformed sib scores that are uncorrelated with the average sib score. Then for the $i$th group, Srivastava (1984) alternative estimators to the maximum likelihood estimates are given by

$$
\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij1}, \quad \hat{\sigma}_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij1} - \bar{y}_{i1})^2 + \bar{a}_i \hat{\gamma}_i^2,
$$

and

$$
\hat{\rho}_i = 1 - \frac{\hat{\gamma}_i^2}{\hat{\sigma}_i^2},
$$

where

$$
\hat{\gamma}_i^2 = \left\{ \sum_{j=1}^{n_i} \sum_{k=2}^{m_{ij}} y_{ijk}^2 \right\} / \left\{ \sum_{j=1}^{n_i} (m_{ij} - 1) \right\},
$$

and

$$
\bar{y}_{i1} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij1}, \quad \bar{a}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} a_{ij}, \quad a_{ij} = 1 - \frac{1}{m_{ij}}.
$$

Note that these estimators are independent of the transformation used except for the estimator of $\mu_i$. However, one may simply choose to use the modified Helmert matrix given earlier.

Srivastava and Katapa (1986) not only the $\hat{\rho}_i$, which is a form of Smith’s (1957) non-iterative uniform weight estimator, has very high efficiency, in that its asymptotic variance is close to that of the maximum likelihood estimate, only if the intraclass correlation is greater than about 0.3. Noting this and the fact that another non-iterative family-size weighted estimator has very high efficiency, in that its asymptotic variance is close to that of the maximum likelihood estimate, only if the intraclass correlation is less than 0.3, Srivastava (1993) suggested an estimator which is an efficient combination of two non-iterative estimators proposed by Smith (1957) and showed that this combination estimator has better efficiency than either one of them. Also see Keen (1993).

For the $i$th population, the two estimators considered by Srivastava (1993) are

$$
\hat{\rho}_{wi} = 1 - \frac{b_{wi} \sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} (x_{ijk} - \bar{x}_{ij})^2}{(m_i - n_i) \sum_{j=1}^{n_i} m_{ij} (\bar{x}_{ij} - \bar{x}_{wi})^2 + a_{wi} \sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} (x_{ijk} - \bar{x}_{ij})^2},
$$

where $a_{wi} = b_{wi} - (n_i - 1), \ b_{wi} = m_i - m_i^{-1} \sum_{j=1}^{n_i} m_{ij}^2, \ m_i = \sum_{j=1}^{n_i} m_{ij}, \ \bar{x}_{wi} = m_i^{-1} \sum_{j=1}^{n_i} m_{ij} \bar{x}_{ij}, \ \bar{x}_{ij} = m_i^{-1} \sum_{k=1}^{m_{ij}} x_{ijk}, \ \text{and}

$$
\hat{\rho}_{ui} = 1 - \frac{b_{ui} \sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} (x_{ijk} - \bar{x}_{ij})^2}{(m_i - n_i) \sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} (x_{ijk} - \bar{x}_{ui})^2 + a_{ui} \sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} (x_{ijk} - \bar{x}_{ij})^2},
$$

where $a_{ui} = (n_i - 1) n_i^{-1} \sum_{j=1}^{n_i} (1 - m_{ij}), \ b_{ui} = (n_i - 1)$.

Srivastava (1993) proposed an improved estimator of $\rho_i$ based on an efficient combination of $\hat{\rho}_{wi}$ and $\hat{\rho}_{ui}$ as

$$
\hat{\rho}_{si} = \frac{\hat{\rho}_{wi}}{1 + \hat{\rho}_{wi} - \hat{\rho}_{ui}},
$$

and has shown that this estimator possesses increased asymptotic relative efficiency over either $\hat{\rho}_{ui}$ or $\hat{\rho}_{wi}$.

3.1. Estimation of the common intraclass correlation

In the case of two populations, Young and Bhandary (1998) estimated the common sib–sib correlation assuming an equal variances for the two populations. The common estimator was obtained by using Srivastava’s (1984) formula (5),
but using data from both the populations. We note that there can be several choices of the common estimate of sib–sib correlation when population variances are different. One simple estimator of the common intraclass correlation under the null hypothesis \( H_0 : \rho_1 = \cdots = \rho_g (=\rho) \) using Srivastava’s (1984) formula (5) is

\[
\hat{\rho} = 1 - \frac{1}{g} \sum_{i=1}^{g} \frac{\hat{\gamma}_i^2}{\hat{\sigma}_i^2}. \tag{8}
\]

This of course is the simple average of the estimators from each population. Another estimator of the common intraclass correlation using Srivastava’s (1993) combination estimator can be suggested as follows. For \( i = 1, \ldots, g \) suppose we write \( \hat{\rho}_{w1} = 1 - \frac{w_{1i}}{w_{2i}} \), where \( w_{1i} \) and \( w_{2i} \), respectively, are the numerator and denominator expressions on the right-hand side of (6) and similarly \( \hat{\rho}_{ui} = 1 - \frac{u_{1i}}{u_{2i}} \), where \( u_{1i} \) and \( u_{2i} \), respectively, the numerator and denominator expressions on the right-hand side of (7).

Let \( \hat{\rho}_w = 1 - (1/g) \sum_{i=1}^{g} w_{1i}/w_{2i} \) and \( \hat{\rho}_u = 1 - (1/g) \sum_{i=1}^{g} u_{1i}/u_{2i} \). Then we suggest

\[
\hat{\rho}_S = \frac{\hat{\rho}_w - \hat{\rho}_u}{1 + \hat{\rho}_w - \hat{\rho}_u}, \tag{9}
\]

as an estimator of the common intraclass correlation. Our initial simulations indicated that \( \hat{\rho}_S \) is a better choice; hence for our subsequent analysis we only considered \( \hat{\rho}_S \). It can be shown that \( \hat{\rho}_S \) is a weighted average of \( \hat{\rho}_{SI} \). However, weights are random, in that, they involve \( \hat{\rho}_{ui} \) and \( \hat{\rho}_{wi} \). Due to this fact, writing an expression for the standard error of \( \hat{\rho}_S \) is difficult, but an expression to determine approximate standard error can be provided. Since this expression is very complicated we suggest using bootstrap method for estimating the standard error and that is what we have adopted in our illustrated example that follows in Section 5.

Knowing that the derivation of the maximum likelihood estimates involves numerical optimization, as in Young and Bhandary (1998), one may suggest modification to the three likelihood-based tests which is obtained by substituting the estimators of \( \mu_i, \sigma_i^2, \rho_i \), specified in (5) and of \( \rho \) in (9) for the maximum likelihood estimators in the expressions of the three likelihood-based tests, as given in (1)–(4). However, this practice may lead to difficulties unless of course, one uses the distribution of the modified test statistics as in Konishi and Gupta (1989). For example, we found out during our simulation studies that thus modified (negative two times the) likelihood ratio was negative, sometimes, for as much as 25% of the simulated data sets. However, as observed in Young and Bhandary (1998), since Srivastava’s estimators are CAN estimators, for really large samples the modified likelihood ratio may approximate the true likelihood ratio.

4. Alternative tests

Young and Bhandary (1998) suggested two asymptotic tests based on normal distribution for testing the equality of two intraclass correlations when the variances of the two populations are equal. In the same spirit here we suggest two asymptotic tests for testing \( H_0 : \rho_1 = \cdots = \rho_g \), based on chi-square distribution. The proposed tests essentially are based on the Srivastava’s combined estimator and its asymptotic variance. The two proposed tests are different only in the way the variances are computed.

Srivastava (1993) provided an expression for the asymptotic variance of his combined estimator. Using his formula the asymptotic variance of \( \hat{\rho}_{SI} \) is given by

\[
AV(\hat{\rho}_{SI}) = v(\rho_i) = 2(1 - \rho_i)^2 \left\{ \frac{(1 - \rho_i)^2}{b_{ui}^2} tr(A_iD_{ii}^2) + \frac{\rho_i^2}{b_{ui}^2} tr(B_iD_{ii}^2) \right\}^2 + \frac{2\rho_i(1 - \rho_i)}{b_{ui}b_{ui}} tr(A_iD_{ii}^2B_iD_{ii}^2) + \frac{[b_{ui} - a_{ui}(1 - \rho_i) - (a_{ui}b_{ui}/b_{ui} - a_{ui})(1 - \rho_i)^2]^2}{b_{ui}^2(m_i - n_i)},
\]

where

\[
D_{ii} = \text{diag}(\eta_{i1}^2, \ldots, \eta_{ni}^2), \quad \eta_{ij}^2 = 1 - (1 - \rho_i)a_{ij}, \quad a_{ij} = 1 - m_{ij}^{-1},
\]

\[
A_i = D_{oi} - m_i^{-1} \omega_i \omega_i', \quad \omega_i = (m_{i1}, \ldots, m_{ni})', \quad B_i = I_{n_i} - n_i^{-1} n_{i} I_{n_i}.'
\]
Note that the asymptotic variance of $\hat{\rho}_{S_i}$, that is, $v(\hat{\rho}_i)$ is a function only of $\rho_i$. Let

$$V_0 = \text{diag}(v(\hat{\rho}_S), \ldots, v(\hat{\rho}_S)) \quad \text{and} \quad V_1 = \text{diag}(v(\hat{\rho}_{S1}), \ldots, v(\hat{\rho}_{Sg})),$$

that is, $V_0$ and $V_1$, respectively, are the estimated variance–covariance matrices of $\hat{\rho}_S = (\hat{\rho}_{S1}, \ldots, \hat{\rho}_{Sg})'$ under the restrictions of null hypothesis and under no restrictions.

We propose the following two tests for testing $H_0: \rho_1 = \cdots = \rho_g$.

Let $T_0$ and $T_1$ be the test statistics defined as

$$T_0 = (C\hat{\rho}_S)'[CV_0C']^{-1}(C\hat{\rho}_S), \quad (10)$$

and

$$T_1 = (C\hat{\rho}_S)'[CV_1C']^{-1}(C\hat{\rho}_S), \quad (11)$$

where the matrix $C$ is as given in (3). Under $H_0$ both, $T_0$ and $T_1$, have asymptotic $\chi^2$ distribution with $(g - 1)$ degrees of freedom.

5. Analysis of Galton’s data

Recently, Hanley (2004) worked with family data on human stature obtained directly from Galton’s note books (cf. Galton, 1886, 1889). The data consist of heights of 205 families with the number of children ranging from 1 to 15. Over all, there were 962 children, 486 of them were sons and the remaining 476 were daughters. However, only 934 children had numerical values. For more details on this data set see Hanley (2004).

For an illustration of our procedures, we divide Galton’s data set into two groups; the first group contains 102 families and the second group contains the remaining 103 families. From the first group we consider data on only daughters and from the second we consider data on only sons. For the first group, the pairs: (the number of daughters, the number of families with those many daughters) are (1, 25), (2, 21), (3, 12), (4, 10), (5, 5), (6, 4), (7, 1), (8, 1), and (9, 1). That is, there are 25 families with one daughter, 21 families with two daughters and so on. Similarly these pairs for the second group from where only sons are selected are (1, 10), (2, 28), (3, 22), (4, 11), (5, 4), and (6, 4). If $\rho_1$ is the correlation between the daughters from the first group and $\rho_2$ is the correlation between the sons from the second group then our interest is to test the null hypothesis $\rho_1 = \rho_2 (= \rho.)$ The maximum likelihood estimates (and their standard errors) of $\rho_1$, $\rho_2$, and $\rho$, respectively, are $\hat{\rho}_1 = 0.2938$ (0.0802), $\hat{\rho}_2 = 0.2023$ (0.0770), and $\hat{\rho} = 0.2489$ (0.0638*). Srivastava’s estimators are $\hat{\rho}_1 = 0.3080$ (0.0842), $\hat{\rho}_2 = 0.2016$ (0.0805), and $\hat{\rho} = 0.2545$ (0.0659*) (* indicates standard error obtained using bootstrap method).

Next, we compute the test statistics and the corresponding $P$-values for all the five tests discussed earlier. The results are summarized in the table below.

<table>
<thead>
<tr>
<th>Test: Statistic</th>
<th>LRT</th>
<th>Score</th>
<th>Wald</th>
<th>$T_0$</th>
<th>$T_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6895</td>
<td>0.7072</td>
<td>0.6775</td>
<td>0.8258</td>
<td>0.8353</td>
</tr>
<tr>
<td></td>
<td>0.4063</td>
<td>0.4105</td>
<td>0.3635</td>
<td>0.4004</td>
<td>0.3607</td>
</tr>
</tbody>
</table>

Clearly, all tests failed to reject $H_0$ at $\alpha = 0.05$ significant level.

We want to point out that our model, having assumed different means and variances for different groups, does account for any differences in different characteristics of the groups. Since the correlation coefficients are location and scale invariant the test provided here indeed tests for the differences in the correlation coefficients from different groups.

6. Simulation experiment and results

Asymptotically, all the tests considered here have chi-square distribution with $g$ degrees of freedom and they are expected to behave similarly for the large sample sizes. However, to assess the performance of all the tests for reasonably small samples, we resolved to the following simulation experiment. Multivariate normal random vectors, for two groups
(i.e. \( g = 2 \)) with \( n_1 = 25 \) samples from the first group and \( n_2 = 35 \) samples from the second group and for three groups (i.e. \( g = 3 \)) with \( n_1 = 25 \) samples from the first group, \( n_2 = 30 \) samples from the second group, and \( n_3 = 35 \) from the third group, are simulated. Following previous simulation experiments (Rosner et al., 1977; Srivastava and Keen, 1988; Young and Bhandary, 1998), the family sizes are simulated from a truncated negative binomial distribution with the number of children ranging from 1 through 15 as suggested in Brass (1958). The mean of the distribution is taken as 2.84 and the success probability as 0.483. For each case, we have used 10,000 simulations.

We made the following choices of the parameters: \( \mu_1 = \mu_2 = \mu_3 = 0 \).

**Case 1:** \( g = 2, \sigma_1^2 = 1, \sigma_2^2 = 2 \) and \( \rho_1 \) and \( \rho_2 \) ranged from 0.1 to 0.9 with increments of 0.1.

**Case 2:** \( g = 3, \sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 3, \rho_1 = \rho_2, \) and \( \rho_3 \) ranged from 0.1 to 0.9 with increments of 0.1.

Estimated size and power values are computed for each of these cases. However, in the interest of space, in the following tables and figures we will provide values of sizes and powers only for a few selected set of parameters.

For testing \( H_0: \rho_1 = \rho_2 = \rho_3 = \rho \), Tables 1–3 provide estimated sizes when the assumed level of the test are, \( \alpha = 0.01, 0.05, \) and 0.10, respectively. From the values in Tables 1–3 we can make the following observations: (a) the chi-square distribution does not provide good approximation to Wald’s statistic (\( Wald \)) in the tail regions; (b) the performance of \( T_1 \) is very erratic, in that the estimated sizes for some values of \( \rho \) are greater than the nominal level and are less for some other values of \( \rho \); (c) although, the \( LRT \) performs reasonably well, its estimated sizes are generally slightly larger than the assumed level; (d) the score test (\( Score \)) and \( T_0 \) test seem to achieve the specified nominal level, though the values are generally lower than the nominal level, for most values of the nuisance parameter \( \rho \).

Power values are provided in Tables 4–6 for the nominal level \( \alpha = 0.05 \) and the choice of parameters \( \rho_1 = \rho_2 = 0.2, 0.5, \) and 0.8, respectively, and for different values of \( \rho_3 \). Since only the score and \( T_0 \) tests achieve the assumed level, it is appropriate to compare the powers of only these two tests. From the values in these tables, it is clear that the score test performs consistently better than \( T_0 \), but \( T_0 \) comes in close second. A plot of the power curves for these tests given in Fig. 1 provides a better view of how close the performances of these two tests are.
Table 3
Sizes, $\alpha = 0.10$, $H_0: \rho_1 = \rho_2 = \rho_3 = \rho$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$T_{0}$</th>
<th>$T_{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.0738</td>
<td>0.0504</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1085</td>
<td>0.1335</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1131</td>
<td>0.1632</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1153</td>
<td>0.1571</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1115</td>
<td>0.1436</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1108</td>
<td>0.1259</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1083</td>
<td>0.1089</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1103</td>
<td>0.0911</td>
</tr>
<tr>
<td>0.9</td>
<td>0.1058</td>
<td>0.0742</td>
</tr>
</tbody>
</table>

Table 4
Rejection proportions, $H_0: \rho_1 = \rho_2 = \rho_3; \rho_1 = \rho_2 = 0.2$

<table>
<thead>
<tr>
<th>$\rho_3$</th>
<th>$T_{0}$</th>
<th>$T_{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.0885</td>
<td>0.0875</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0554</td>
<td>0.0720</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0983</td>
<td>0.1388</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2431</td>
<td>0.3098</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4903</td>
<td>0.5671</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7651</td>
<td>0.8176</td>
</tr>
<tr>
<td>0.7</td>
<td>0.9487</td>
<td>0.9632</td>
</tr>
<tr>
<td>0.8</td>
<td>0.9968</td>
<td>0.9976</td>
</tr>
<tr>
<td>0.9</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5
Rejection proportions, $H_0: \rho_1 = \rho_2 = \rho_3; \rho_1 = \rho_2 = 0.5$

<table>
<thead>
<tr>
<th>$\rho_3$</th>
<th>$T_{0}$</th>
<th>$T_{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.7312</td>
<td>0.7982</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4634</td>
<td>0.5443</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2329</td>
<td>0.2952</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1021</td>
<td>0.1398</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0620</td>
<td>0.0901</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1180</td>
<td>0.1471</td>
</tr>
<tr>
<td>0.7</td>
<td>0.3549</td>
<td>0.3871</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7898</td>
<td>0.8017</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9966</td>
<td>0.9966</td>
</tr>
</tbody>
</table>

Table 6
Rejection proportions, $H_0: \rho_1 = \rho_2 = \rho_3; \rho_1 = \rho_2 = 0.8$

<table>
<thead>
<tr>
<th>$\rho_3$</th>
<th>$T_{0}$</th>
<th>$T_{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.9998</td>
<td>0.9999</td>
</tr>
<tr>
<td>0.2</td>
<td>0.9975</td>
<td>0.9976</td>
</tr>
<tr>
<td>0.3</td>
<td>0.9806</td>
<td>0.9800</td>
</tr>
<tr>
<td>0.4</td>
<td>0.9250</td>
<td>0.9152</td>
</tr>
<tr>
<td>0.5</td>
<td>0.7631</td>
<td>0.7363</td>
</tr>
<tr>
<td>0.6</td>
<td>0.4779</td>
<td>0.4411</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1849</td>
<td>0.1661</td>
</tr>
<tr>
<td>0.8</td>
<td>0.0547</td>
<td>0.0420</td>
</tr>
<tr>
<td>0.9</td>
<td>0.4343</td>
<td>0.3667</td>
</tr>
</tbody>
</table>
Next, for testing $H_0 : \rho_1 = \rho_2$, with $\alpha = 0.05$, we provide the values of power in Tables 7–9, respectively, for $\rho_1 = 0.2$, 0.5, and 0.8. The power curves for the score and $T_0$ tests are provided in Fig. 2. From the values in the tables and from the power curves, we see that the performance of $T_0$ test is some times better than that of score test, but overall performances of these two tests are essentially indistinguishable.
Thus, based on our limited power study of these tests, we recommend using score test or $T_0$ test. Since $T_0$ can be computed directly without going through numerical evaluations, we may in fact, prefer this test in practice.

It is fairly well-known that the score test is locally most powerful (Mukherjee, 1993) and tends to do better than its competitors, the Wald’s and the LRTs. In the definition of $T_0$ (see (10)), the variance–covariance matrix $V_0$ is estimated under the null hypothesis $H_0: \rho_1 = \cdots = \rho_g$. This process of evaluating the variance–covariance matrix under the null hypothesis is, at least in principle, similar to the construction of score test. This perhaps is a reason why $T_0$ performs well and its performance is similar to that of score test. However, we must emphasize that these conclusions are based on the limited simulation study that we have undertaken here.

Table 7
Rejection proportions, $H_0: \rho_1 = \rho_2; \rho_1 = 0.2$

<table>
<thead>
<tr>
<th>$\rho_2$</th>
<th>LRT</th>
<th>Wald</th>
<th>Score</th>
<th>$T_0$</th>
<th>$T_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.0848</td>
<td>0.0738</td>
<td>0.0943</td>
<td>0.0810</td>
<td>0.0936</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0537</td>
<td>0.0628</td>
<td>0.0491</td>
<td>0.0467</td>
<td>0.0585</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1084</td>
<td>0.1408</td>
<td>0.0820</td>
<td>0.0855</td>
<td>0.1092</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2359</td>
<td>0.2909</td>
<td>0.1923</td>
<td>0.1941</td>
<td>0.2322</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4461</td>
<td>0.5069</td>
<td>0.3904</td>
<td>0.3973</td>
<td>0.4358</td>
</tr>
<tr>
<td>0.6</td>
<td>0.6920</td>
<td>0.7295</td>
<td>0.6422</td>
<td>0.6499</td>
<td>0.6747</td>
</tr>
<tr>
<td>0.7</td>
<td>0.8998</td>
<td>0.9101</td>
<td>0.8730</td>
<td>0.8771</td>
<td>0.8829</td>
</tr>
<tr>
<td>0.8</td>
<td>0.9906</td>
<td>0.9908</td>
<td>0.9861</td>
<td>0.9886</td>
<td>0.9873</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9990</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

Table 8
Rejection proportions, $H_0: \rho_1 = \rho_2; \rho_1 = 0.5$

<table>
<thead>
<tr>
<th>$\rho_2$</th>
<th>LRT</th>
<th>Wald</th>
<th>Score</th>
<th>$T_0$</th>
<th>$T_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.6540</td>
<td>0.6795</td>
<td>0.6385</td>
<td>0.6037</td>
<td>0.6478</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4461</td>
<td>0.5069</td>
<td>0.3904</td>
<td>0.3973</td>
<td>0.4358</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2194</td>
<td>0.2487</td>
<td>0.2178</td>
<td>0.1930</td>
<td>0.2330</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1008</td>
<td>0.1179</td>
<td>0.0959</td>
<td>0.0866</td>
<td>0.1090</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0565</td>
<td>0.0687</td>
<td>0.0494</td>
<td>0.0483</td>
<td>0.0583</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1245</td>
<td>0.1357</td>
<td>0.1056</td>
<td>0.1090</td>
<td>0.1127</td>
</tr>
<tr>
<td>0.7</td>
<td>0.3284</td>
<td>0.3159</td>
<td>0.2941</td>
<td>0.2963</td>
<td>0.2821</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7153</td>
<td>0.6786</td>
<td>0.6790</td>
<td>0.6803</td>
<td>0.6429</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9855</td>
<td>0.9761</td>
<td>0.9811</td>
<td>0.9814</td>
<td>0.9705</td>
</tr>
</tbody>
</table>

Table 9
Rejection proportions, $H_0: \rho_1 = \rho_2; \rho_1 = 0.8$

<table>
<thead>
<tr>
<th>$\rho_2$</th>
<th>LRT</th>
<th>Wald</th>
<th>Score</th>
<th>$T_0$</th>
<th>$T_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.9975</td>
<td>0.9980</td>
<td>0.9962</td>
<td>0.9954</td>
<td>0.9974</td>
</tr>
<tr>
<td>0.2</td>
<td>0.9906</td>
<td>0.9908</td>
<td>0.9861</td>
<td>0.9886</td>
<td>0.9873</td>
</tr>
<tr>
<td>0.3</td>
<td>0.9538</td>
<td>0.9623</td>
<td>0.9495</td>
<td>0.9418</td>
<td>0.9565</td>
</tr>
<tr>
<td>0.4</td>
<td>0.8724</td>
<td>0.8832</td>
<td>0.8699</td>
<td>0.8481</td>
<td>0.8755</td>
</tr>
<tr>
<td>0.5</td>
<td>0.7153</td>
<td>0.6786</td>
<td>0.6790</td>
<td>0.6803</td>
<td>0.6429</td>
</tr>
<tr>
<td>0.6</td>
<td>0.4544</td>
<td>0.4585</td>
<td>0.4530</td>
<td>0.4107</td>
<td>0.4500</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1818</td>
<td>0.1771</td>
<td>0.1823</td>
<td>0.1547</td>
<td>0.1728</td>
</tr>
<tr>
<td>0.8</td>
<td>0.0572</td>
<td>0.0440</td>
<td>0.0528</td>
<td>0.0456</td>
<td>0.0390</td>
</tr>
<tr>
<td>0.9</td>
<td>0.3952</td>
<td>0.2794</td>
<td>0.3639</td>
<td>0.3524</td>
<td>0.2542</td>
</tr>
</tbody>
</table>
Fig. 2. Power estimation of sib–sib correlation using $\alpha = 0.05$ for testing $H_0: \rho_1 = \rho_2 = \rho$.

Acknowledgments

We want to thank the referees whose many suggestions helped improve the paper considerably.

References