Journal of Modern Applied Statistical Methods

Volume 13 | Issue 1 Article 4

5-1-2014

An Alternative Test for the Equality of Intraclass Correlation Coefficients under Unequal Family Sizes for Several Populations

Madhusudan Bhandary
Columbus State University, bhandary_madhusudan@colstate.edu

Koji Fujiwara
North Dakota State University, koji.fujiwara@ndsu.edu

Journal of Modern Applied Statistical Methods
May 2014, Vol. 13, No. 1, 36-54.

ISSN 1538 − 9472

An Alternative Test for the Equality of Intraclass Correlation Coefficients under Unequal Family Sizes for Several Populations

Madhusudan Bhandary
Columbus State University
Columbus, GA

Koji Fujiwara
North Dakota State University
Fargo, ND

An alternative test for the equality of several intraclass correlation coefficients under unequal family sizes based on several independent multinormal samples is proposed. It was found that the alternative test consistently and reliably produced results superior to those of Likelihood ratio test (LRT) proposed by Bhandary and Alam (2000) and F_{max} test proposed by Bhandary and Fujiwara (2006) in terms of power for various combinations of intraclass correlation coefficient values and also the alternative test stays closer to the significance level under null hypothesis compared to the Likelihood ratio test and F_{max} test. This alternative test is computationally very simple and also can be used for both small sample and large sample situations. An example with real life data is presented.

Keywords: Likelihood ratio test, F_{max}-test, Alternative test, intraclass correlation coefficient

Introduction

It is sometimes necessary to estimate the correlation coefficient between blood pressures of children on the basis of measurements taken on p children in each of n families. The p measurements on a family provide p (p − 1) pairs of observations (x,y), x being the blood pressure of one child and y that of another. From the n families a total of np (p − 1) pairs are generated from which a correlation coefficient is computed in the ordinary way.

Dr. Bhandary is a Professor in the Department of Mathematics. Email him at: bhandary_madhusudan@colstate.edu. Koji Fujiwara is a graduate student in the Department of Statistics. Email him at: koji.fujiwara@ndsu.edu.

The correlation coefficient thus computed is called intraclass correlation coefficient. It is important to have statistical inference concerning intraclass correlation, because it provides information regarding blood pressure, cholesterol etc. in a family within some race in the world.

The intraclass correlation coefficient ρ has a wide variety of applications. It can be used to measure the degree of intra-family resemblance with respect to characteristics such as blood pressure, cholesterol, weight, height, stature, lung capacity, etc.

Statistical inference concerning ρ based on a single multinormal sample has been studied by several authors (Scheffe, 1959; Rao, 1973; Rosner, et al., 1977, 1979; Donner and Bull, 1983; Srivastava, 1984; Konishi, 1985; Gokhale and SenGupta, 1986; SenGupta, 1988; Velu and Rao, 1990).

For a two sample problem, Donner and Bull (1983) discussed the likelihood ratio test for testing the equality of two intraclass correlation coefficients based on two independent multinormal samples under equal family sizes. Konishi and Gupta (1987) proposed a modified likelihood ratio test and derived its asymptotic null distribution. They also discussed another test procedure based on a modification of Fisher's Z-transformation following Konishi (1985).

For a several sample problem, Huang and Sinha (1993) considered an optimum invariant test for the equality of intraclass correlation coefficients under equal family sizes for more than two intraclass correlation coefficients based on independent samples from several multinormal distributions.

For unequal family sizes, Young and Bhandary (1998) proposed Likelihood ratio test, large sample Z-test and large sample Z^*-test for the equality of two intraclass correlation coefficients based on two independent multinormal samples.

For several populations and unequal family sizes, Bhandary and Alam (2000) proposed Likelihood ratio test and large sample ANOVA test for the equality of several intraclass correlation coefficients based on several independent multinormal samples. Bhandary and Fujiwara (2006) proposed $F_{\max}$ test for the equality of several intraclass correlation coefficients under unequal family sizes. Donner and Zou (2002) proposed asymptotic test for the equality of dependent intraclass correlation coefficients under unequal family sizes.

An alternative test for the equality of several intraclass correlation coefficients is considered based on several independent multinormal samples under unequal family sizes.

A conditional analysis is carried out here, assuming family sizes fixed though unequal.

It could be of interest to see whether blood pressure or cholesterol or lung capacity, etc., among families in Caucasian, Asian, Hispanic or African races, etc., differ or not; therefore a small sample test for the equality of intraclass correlation coefficients under unequal family sizes has been developed.

Also, an alternative test is proposed for the equality of intraclass correlation coefficients under unequal family sizes, which is computationally very simple. A brief discussion of likelihood ratio test proposed by Bhandary and Alam (2000) and $F_{\max}$ test proposed by Bhandary and Fujiwara (2006) are provided.

These tests are compared in the section titled *Simulation Results*, using simulation technique. It is found on the basis of simulation study that the alternative test consistently and reliably produced results superior to those of Likelihood ratio test and $F_{\max}$ test in terms of power for various combination of intraclass correlation coefficient values and also the alternative test stays closer to the significance level under null hypothesis compared to the Likelihood ratio test and $F_{\max}$ test.

An example with real life data is given in the section titled *Example With Real Life Data*.

Tests of $H_0 : \rho_1 = \rho_2 = \rho_3$ Versus H_1 : *NOT* H_0

Likelihood Ratio Test

Let $\underset{\sim}{X}_i = (x_{i1}, x_{i2}, \ldots, x_{ip_i})'$ be a $p_i x1$ vector of observations from the i^{th} family; $i = 1, 2, \ldots, k$. The structure of mean vector and the covariance matrix for the familial data is given by the following (Rao, 1973):

$$\underset{\sim}{\mu}_i = \mu \underset{\sim}{1}_i \text{ and } \sum_{p_i x p_i} {}_i = \sigma^2 \begin{pmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \ldots & \rho \\ \ldots & \ldots & \ldots & \ldots \\ \rho & \rho & \ldots & 1 \end{pmatrix} \tag{1}$$

where $\underset{\sim}{1}_i$ is a $p_i x1$ vector of 1's, $\mu(-\infty < \mu < \infty)$ is the common mean and $\sigma^2(\sigma^2 > 0)$ is the common variance of members of the family and ρ, which is called the intraclass correlation coefficient, is the coefficient of correlation among the members of the family and $\max_{1 \le i \le k} \left(-\frac{1}{p_i - 1} \right) \le \rho \le 1$.

It is assumed that $\underset{\sim}{x}_i \sim N_{p_i}(\underset{\sim}{\mu}_i, \Sigma_i); i = 1,...,k$, where N_{p_i} represents p_i-variate normal distribution and $\underset{\sim}{\mu}_i, \Sigma_i$'s are defined in (1).

Let

$$\underset{\sim}{u}_i = (u_{i1}, u_{i2}, ..., u_{ip_i})' = Q \underset{\sim}{X}_i \tag{2}$$

where Q is an orthogonal matrix.

Under the orthogonal transformation (2), it can be seen that $\underset{\sim}{u}_i \sim N_{p_i}(\underset{\sim}{\mu}_i^*, \Sigma_i^*); i = 1,...,k$ where $\underset{p_i x1}{\underset{\sim}{\mu}_i^*} = (\mu, 0, 0, ..., 0)'$ and

$$\sum\nolimits_i^* = \sigma^2 \begin{pmatrix} \eta_i & 0 & \dots & 0 \\ 0 & 1-\rho & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1-\rho \end{pmatrix} \text{ and } \eta_i = p_i^{-1}\left\{1 + \left(p_i - 1\right)\rho\right\}.$$

The transformation used on the data from $\underset{\sim}{x}$ to $\underset{\sim}{u}$ above is independent of ρ. One can use Helmert's orthogonal transformation.

Srivastava (1984) gives estimator of ρ and σ^2 under unequal family sizes which are good substitute for the maximum likelihood estimator and are given by the following:

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

$$\hat{\sigma}^2 = (k-1)^{-1}\sum_{i=1}^{k}(u_{i1} - \hat{\mu})^2 + k^{-1}\hat{\gamma}^2(\sum_{i=1}^{k} a_i)$$

$$\text{where } \hat{\gamma}^2 = \frac{\sum_{i=1}^{k}\sum_{r=2}^{p_i} u_{ir}^2}{\sum_{i=1}^{k}(p_i - 1)} \tag{3}$$

$$\hat{\mu} = k^{-1}\sum_{i=1}^{k} u_{i1}$$

$$\text{and } a_i = 1 - p_i^{-1}.$$

Now, consider the three sample problem with k_1, k_2 and k_3 families from each population.

Let $\underset{\sim}{x_i} = (x_{i1}, x_{i2}, ..., x_{ip_i})'$ be a $p_i x1$ vector of observations from i^{th} family; $i = 1,...,k_1$

$$\text{and } \underset{\sim}{x_i} \sim N_{p_i}(\underset{\sim}{\mu_{1i}}, \Sigma_{1i}), \text{ where } \underset{\sim}{\mu_{1i}} = \mu_1 \underset{\sim}{1_i}, \Sigma_{1i} = \sigma_1^2 \begin{pmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_1 \\ \dots & \dots & \dots & \dots \\ \rho_1 & \rho_1 & \dots & 1 \end{pmatrix} \quad (4)$$

and $\max_{1 \le i \le k_1} \left(-\frac{1}{p_i - 1} \right) \le \rho_1 \le 1$

Let $\underset{\sim}{y_j} = (y_{j1}, y_{j2}, ..., y_{jq_j})'$ be a $q_j x1$ vector of observations from j^{th} family in the second population; $j = 1, \dots, k_2$

$$\text{and } \underset{\sim}{y_j} \sim N_{q_j}(\underset{\sim}{\mu_{2j}}, \Sigma_{2j})$$

$$\text{where } \underset{\sim}{\mu_{2j}} = \mu_2 \underset{\sim}{1_j}, \Sigma_{2j} = \sigma_2^2 \begin{pmatrix} 1 & \rho_2 & \dots & \rho_2 \\ \rho_2 & 1 & \dots & \rho_2 \\ \dots & \dots & \dots & \dots \\ \rho_2 & \rho_2 & \dots & 1 \end{pmatrix} \quad (5)$$

$$\text{and } \max_{1 \le j \le k_2} \left(-\frac{1}{q_j - 1} \right) \le \rho_2 \le 1$$

Let $\underset{\sim}{z_l} = (z_{l1}, z_{l2}, ..., z_{lr_l})'$ be a $r_l x1$ vector of observations from l^{th} family; $l = 1, 2, ..., k_3$

$$\text{and } \underset{\sim}{z_l} \sim N_{r_l}(\underset{\sim}{\mu_{3l}}, \Sigma_{3l}), \text{ where } \underset{\sim}{\mu_{3l}} = \mu_3 \underset{\sim}{1_l}, \Sigma_{3l} = \sigma_3^2 \begin{pmatrix} 1 & \rho_3 & \dots & \rho_3 \\ \rho_3 & 1 & \dots & \rho_3 \\ \dots & \dots & \dots & \dots \\ \rho_3 & \rho_3 & \dots & 1 \end{pmatrix} \quad (6)$$

and $\max_{1 \le l \le k_3}\left(-\frac{1}{r_l - 1}\right) \le \rho_3 \le 1$.

Using orthogonal transformation, the data vector can be transformed from $\underset{\sim}{x}_i$ to $\underset{\sim}{u}_i$, $\underset{\sim}{y}_j$ to $\underset{\sim}{v}_j$ and $\underset{\sim}{z}_l$ to $\underset{\sim}{w}_l$ as follows:

$$\underset{\sim}{u}_i = (u_{i1}, u_{i2}, ..., u_{ip_i})' \sim N_{p_i}(\underset{\sim}{\mu}^*_{1i}, \Sigma^*_{1i}); i = 1, ..., k_1$$

and $$\underset{\sim}{v}_j = (v_{j1}, v_{j2}, ..., v_{jq_j})' \sim N_{q_j}(\underset{\sim}{\mu}^*_{2j}, \Sigma^*_{2j}); j = 1, ..., k_2$$

and $$\underset{\sim}{w}_l = (w_{l1}, w_{l2}, ..., w_{lr_l})' \sim N_{r_l}(\underset{\sim}{\mu}^*_{3l}, \Sigma^*_{3l}); l = 1, ..., k_3$$

where, $$\underset{\sim}{\mu}^*_{1i} = (\mu_1, 0, 0, ..., 0)', \Sigma^*_{1i} = \sigma_1^2 \begin{pmatrix} \eta_i & 0 & ... & 0 \\ 0 & 1-\rho_1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1-\rho_1 \end{pmatrix}$$

$$\eta_i = p_i^{-1}\{1 + (p_i - 1)\rho_1\} \tag{7}$$

$$\underset{\sim}{\mu}^*_{2j} = (\mu_2, 0, 0, ..., 0)', \Sigma^*_{2j} = \sigma_2^2 \begin{pmatrix} \xi_j & 0 & ... & 0 \\ 0 & 1-\rho_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1-\rho_2 \end{pmatrix}$$

$$\xi_j = q_j^{-1}\{1 + (q_j - 1)\rho_2\}$$

$$\underset{\sim}{\mu}^*_{3l} = (\mu_3, 0, 0, ..., 0)', \Sigma^*_{3l} = \sigma_3^2 \begin{pmatrix} \varsigma_l & 0 & ... & 0 \\ 0 & 1-\rho_3 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1-\rho_3 \end{pmatrix}$$

and $$\varsigma_l = r_l^{-1}\{1 + (r_l - 1)\rho_3\}$$

The transformations used on the data above from $\underset{\sim}{x}$ to $\underset{\sim}{u}$, $\underset{\sim}{y}$ to $\underset{\sim}{v}$ and $\underset{\sim}{z}$ to $\underset{\sim}{w}$ are independent of ρ_1, ρ_2 and ρ_3. It is assumed that $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$.

Under the above setup, Bhandary and Alam (2000) derived likelihood ratio test statistic for testing $H_0 : \rho_1 = \rho_2 = \rho_3$ Vs. $H_1 : NOT \ H_0$ which is given by the following:

$$\begin{aligned}
-2\log\Lambda \ &= \sum_{i=1}^{k_1} \log\left[p_i^{-1}\left\{1+(p_i-1)\hat{\rho}\right\}\right] + \sum_{i=1}^{k_1}(p_i-1)\log(1-\hat{\rho}) \\
&+\sum_{j=1}^{k_2} \log\left[q_j^{-1}\left\{1+(q_j-1)\hat{\rho}\right\}\right] + \sum_{j=1}^{k_2}(q_j-1)\log(1-\hat{\rho}) \\
&+\sum_{l=1}^{k_3} \log\left[r_l^{-1}\left\{1+(r_l-1)\hat{\rho}\right\}\right] + \sum_{l=1}^{k_3}(r_l-1)\log(1-\hat{\rho}) \\
&+\frac{1}{\hat{\sigma}^2}\left[\sum_{i=1}^{k_1}\left\{ p_i(u_{i1}-\hat{\mu}_1)^2 \Big/ \left[1+(p_i-1)\hat{\rho}\right]\right\} + \sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 / (1-\hat{\rho}) \right. \\
&+\sum_{j=1}^{k_2}\left\{ q_j(v_{j1}-\hat{\mu}_2)^2 \Big/ \left[1+(q_j-1)\hat{\rho}\right]\right\} + \sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2 / (1-\hat{\rho}) \\
&\left.+\sum_{l=1}^{k_3}\left\{ r_l(w_{l1}-\hat{\mu}_3)^2 \Big/ \left[1+(r_l-1)\hat{\rho}\right]\right\} + \sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2 / (1-\hat{\rho}) \right] \\
&-\sum_{i=1}^{k_1} \log\left[p_i^{-1}\left\{1+(p_i-1)\hat{\rho}_1\right\}\right] - \sum_{i=1}^{k_1}(p_i-1)\log(1-\hat{\rho}_1) \\
&-\sum_{j=1}^{k_2} \log\left[q_j^{-1}\left\{1+(q_j-1)\hat{\rho}_2\right\}\right] - \sum_{j=1}^{k_2}(q_j-1)\log(1-\hat{\rho}_2) \\
&-\sum_{l=1}^{k_3} \log\left[r_l^{-1}\left\{1+(r_l-1)\hat{\rho}_3\right\}\right] - \sum_{l=1}^{k_3}(r_l-1)\log(1-\hat{\rho}_3) \\
&-\frac{1}{\hat{\sigma}^2}\left[\sum_{i=1}^{k_1}\left\{ p_i(u_{i1}-\hat{\mu}_1)^2 \Big/ \left[1+(p_i-1)\hat{\rho}_1\right]\right\} + \sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 / (1-\hat{\rho}_1) \right. \\
&+\sum_{j=1}^{k_2}\left\{ q_j(v_{j1}-\hat{\mu}_2)^2 \Big/ \left[1+(q_j-1)\hat{\rho}_2\right]\right\} + \sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2 / (1-\hat{\rho}_2) \\
&\left.+\sum_{l=1}^{k_3}\left\{ r_l(w_{l1}-\hat{\mu}_3)^2 \Big/ \left[1+(r_l-1)\hat{\rho}_3\right]\right\} + \sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2 / (1-\hat{\rho}_3) \right] \qquad (8)
\end{aligned}$$

where, Λ = likelihood ratio test statistic,

$\hat{\rho}$ = estimate of common intraclass correlation coefficients under H_0,

$\hat{\rho}_1$ = estimate of intraclass correlation coefficient from first sample under H_1,

$\hat{\rho}_2$ = estimate of intraclass correlation coefficient from second sample under H_1,

$\hat{\rho}_3$ = estimate of intraclass correlation coefficient from third sample under H_1,

$\hat{\sigma}^2$ = estimate of σ^2

and $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\mu}_3$ are estimates of means from first ,second and third samples respectively.

The estimators $\hat{\rho}, \hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3, \hat{\sigma}^2, \hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\mu}_3$ can be obtained from Srivastava's estimator given by (3).

It is known from asymptotic theory that $-2\log\Lambda$ has an asymptotic chi-square distribution with 2 degrees of freedom.

Bhandary and Alam (2000) also suggested large sample ANOVA test and showed through simulation that likelihood ratio test given by (8) consistently produced results superior to those of the large sample ANOVA test.

The likelihood ratio test given by (8) is computationally complex, and used asymptotically – that is, when family sizes are large (at least 30). But situations may also call for a small sample case. An alternative test is here proposed, which is computationally very simple and can be used for both small sample and large sample situations.

*F*max test

The $F_{\max}$ test is described as follows:

$$F_{\max} = \max\left\{F_1, F_2, F_3, F_4, F_5, F_6\right\} \tag{9}$$

$$\text{where } F_1 = \frac{\sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 \Big/ \left\{\sum_{i=1}^{k_1}(p_i - 1)\right\}}{\sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2 \Big/ \left\{\sum_{j=1}^{k_2}(q_j - 1)\right\}} \tag{10}$$

$$F_2 = \frac{\sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 \Big/ \left\{\sum_{i=1}^{k_1}(p_i - 1)\right\}}{\sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2 \Big/ \left\{\sum_{l=1}^{k_3}(r_l - 1)\right\}} \tag{11}$$

$$F_3 = \frac{\sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2 \Big/ \left\{\sum_{j=1}^{k_2}(q_j - 1)\right\}}{\sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2 \Big/ \left\{\sum_{l=1}^{k_3}(r_l - 1)\right\}} \tag{12}$$

$$F_4 = 1/F_1, F_5 = 1/F_2, \text{ and } F_6 = 1/F_3 \tag{13}$$

It can be shown using (7) that

$$\frac{\sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2}{\sigma^2(1-\rho_1)} \sim \chi_{pp}^2,\ \frac{\sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2}{\sigma^2(1-\rho_2)} \sim \chi_{qq}^2 \text{ and } \frac{\sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2}{\sigma^2(1-\rho_3)} \sim \chi_{rr}^2 \tag{14}$$

where, χ_n^2 denotes chi-square distribution with n degrees of freedom and $pp = \sum_{i=1}^{k_1}(p_i - 1); qq = \sum_{j=1}^{k_2}(q_j - 1); rr = \sum_{l=1}^{k_3}(r_l - 1).$

Therefore, using (14) under H_0, the exact distribution of the F_1 given by (10) is $F_{pp,qq}$. (15)

Similarly, using (14) under H_0, the exact distributions of F_2, F_3, F_4, F_5 and F_6 are $F_{pp,rr}$, $F_{qq,rr}$, $F_{qq,pp}$, $F_{rr,pp}$ and $F_{rr,qq}$ respectively, where F_{n_1,n_2} denotes F-distribution with n_1 and n_2 degrees of freedom respectively. (16)

Journal of Modern Applied Statistical Methods
May 2014, Vol. 13, No. 1, 36-54.

ISSN 1538 − 9472

Hence, using (9), (15) and (16) and using Bonferroni's bound, approximate critical value at α for testing H_0 Vs. H_1 can be proposed as

$$C = \max\left\{ F_{\frac{\alpha}{6};pp,qq}, F_{\frac{\alpha}{6};pp,rr}, F_{\frac{\alpha}{6};qq,rr}, F_{\frac{\alpha}{6};qq,pp}, F_{\frac{\alpha}{6};rr,pp}, F_{\frac{\alpha}{6};rr,qq} \right\} \tag{17}$$

where, $F_{\gamma;a,b}$ is the upper $100\gamma\%$ point of F-distribution with degrees of freedom a and b respectively.

The critical region for testing H_0 Vs. H_1 is proposed as follows:

$$F_{\max} > C \tag{18}$$

where $F_{\max}$ and C are given by (9) and (17), respectively.

The test statistic $F_{\max}$ given by (9) is very simple to compute, and the distributions of F_1, F_2, F_3, F_4, F_5 and F_6 are exact and hence can be used for both small sample and large sample situations.

Alternative test

For the alternative test, the test statistic is described as follows:

$$F_1 = \frac{\sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 \Big/ \left\{\sum_{i=1}^{k_1}(p_i - 1)\right\}}{\left(\sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2 + \sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2\right) \Big/ \left\{\sum_{j=1}^{k_2}(q_j - 1) + \sum_{l=1}^{k_3}(r_l - 1)\right\}}$$

$$F_2 = \frac{\sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2 \Big/ \left\{\sum_{j=1}^{k_2}(q_j - 1)\right\}}{\left(\sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 + \sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2\right) \Big/ \left\{\sum_{i=1}^{k_1}(p_i - 1) + \sum_{l=1}^{k_3}(r_l - 1)\right\}}$$

$$F_3 = \frac{\sum_{l=1}^{k_3}\sum_{t=2}^{r_l} w_{lt}^2 \Big/ \left\{\sum_{l=1}^{k_3}(r_l - 1)\right\}}{(\sum_{i=1}^{k_1}\sum_{r=2}^{p_i} u_{ir}^2 + \sum_{j=1}^{k_2}\sum_{s=2}^{q_j} v_{js}^2) \Big/ \left\{\sum_{i=1}^{k_1}(p_i - 1) + \sum_{j=1}^{k_2}(q_j - 1)\right\}}$$

$$F_4 = 1/F_1, F_5 = 1/F_2 \text{ and } F_6 = 1/F_3 \qquad (19)$$

Using (14), it can be said that under H_0, the exact distribution of the F_1 is $F_{pp,qq+rr}$.

Similarly, under H_0, the exact distributions of F_2, F_3, F_4, F_5 and F_6 are $F_{qq,pp+rr}$, $F_{rr,pp+qq}$, $F_{qq+rr,pp}$, $F_{pp+rr,qq}$ and $F_{pp+qq,rr}$ respectively, where, F_{n_1,n_2} denotes F-distribution with n_1 and n_2 degrees of freedom respectively.

Set the P-values to be the right tail probability of the statistics calculated above such that $P_i = P(X > F_i)$ where F_i's are explained in (19).

Sort the P-values obtained as above in an ascending order and denote them by $P_{(1)}, P_{(2)}, \cdots, P_{(6)}$.

$$\text{Reject } H_0 \text{ if } P_{(i)} < \frac{i}{6}\alpha \text{ for some } i \in \{1, 2, \ldots, 6\}. \qquad (20)$$

In order that H_0 is insignificant, it is required that $P_{(1)} \geq \frac{1}{6}\alpha, P_{(2)} \geq \frac{2}{6}\alpha, ..., P_{(6)} \geq \alpha$. So, if $P_{(i)} < \frac{i}{6}\alpha$ then the test corresponding to $P_{(1)}$ is insignificant, corresponding to $P_{(2)}$ is insignificant, …, corresponding to $P_{(i-1)}$ is insignificant and corresponding to $P_{(i)}$ is significant and the overall test is significant.

Simulation Results

Multivariate normal random vectors were generated using R program in order to evaluate the power of the alternative test as compared to $F_{\max}$ test and the LRT test. Five and thirty vectors of family data were created for each of the three populations. The family size distribution was truncated to maintain the family size at a minimum of 2 siblings and a maximum of 15 siblings. The previous research in simulating family sizes (Rosner et al., 1977; Srivastava and Keen, 1988) determined the parameter setting for FORTRAN IMSL negative binomial subroutine with a mean = 2.86 and a success probability = 0.483. Here, it is set at a mean = 2.86 and a theta = 41.2552.

All parameters were set the same for each population, except the values of ρ_1, ρ_2 and ρ_3 which took various combinations over the range of values from 0.1 to 0.9 at increments of 0.1.

The R program produced estimates of ρ_1, ρ_2 and ρ_3 along with $F_{\max}$ statistic and LRT statistic and the new statistic 10,000 times for each particular combination of population parameters (ρ_1, ρ_2 and ρ_3).

The frequency of rejection of each test at α = 0.05 was noted and the proportion of rejections are noted for a sample combinations of ρ_1, ρ_2 and ρ_3.

The size comparison for the alternative test, $F_{\max}$ test and the LRT test for various combinations of ρ_1, ρ_2 and ρ_3 is also presented.

A few figures are presented of powers estimates as well as size estimates for these tests. On the basis of this study, it was found that the alternative test showed consistently better results in terms of power as well as in size than LRT and $F_{\max}$ test. This alternative test is computationally very simple and also can be used for both small sample and large sample situations. The alternative test stays closer to the significance level under null hypothesis compared to the Likelihood ratio test and $F_{\max}$ test. It is recommend that the alternative test is used in practice.

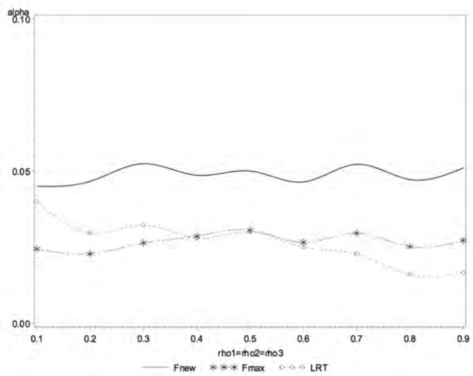

Figure 1. Size Estimates (α = 0.05 and k = 5)

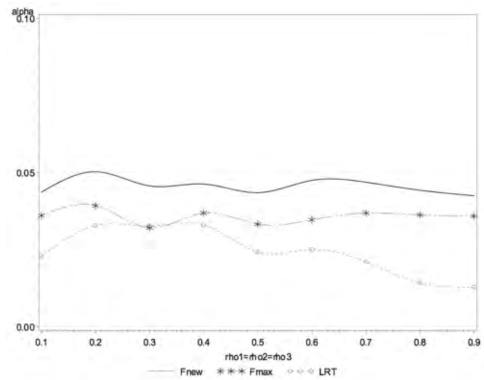

Figure 2. Size Estimates (α = 0.05 and k = 30)

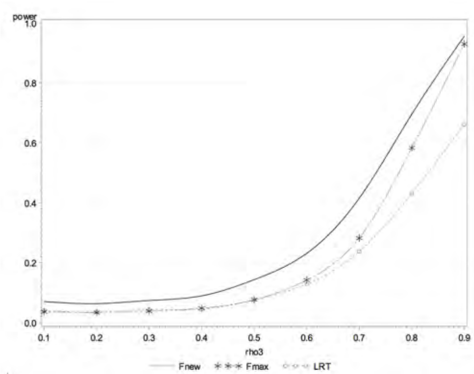

Figure 3. Power Estimates (α = 0.05, k = 5, ρ_1 = 0.1 and ρ_2 = 0.3)

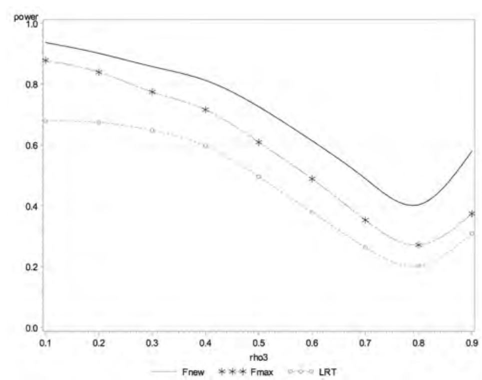

Figure 4. Power Estimates (α = 0.05, k = 5, ρ_1 = 0.7 and ρ_2 = 0.9)

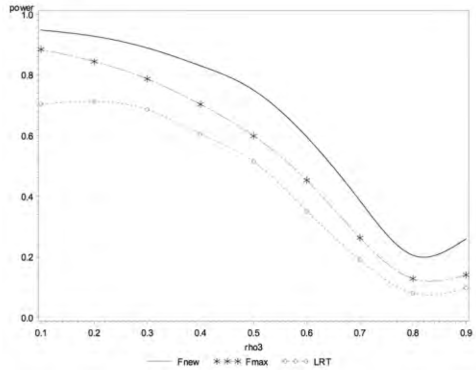

Figure 5. Power Estimates (α = 0.05, k = 5, ρ_1 = 0.9 and ρ_2 = 0.8)

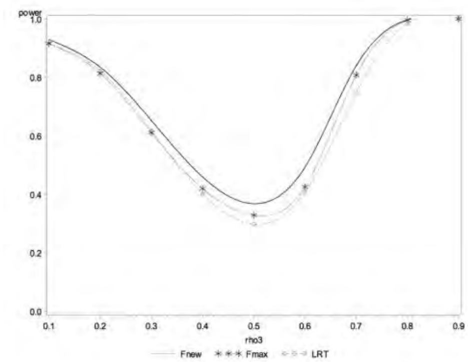

Figure 6. Power Estimates (α = 0.05, k = 30, ρ_1 = 0.4 and ρ_2 = 0.6)

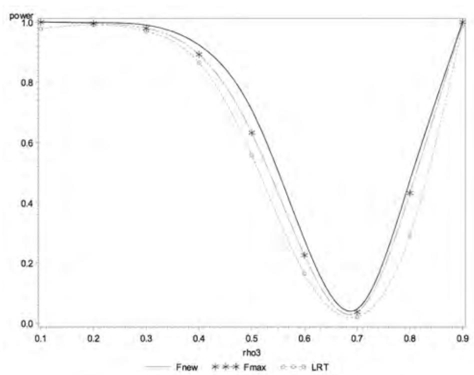

Figure 7. Power Estimates (α = 0.05, k = 30, ρ_1 = 0.7 and ρ_2 = 0.7)

Table 1. Size Estimates (α = 0.05)

	k = 5			k = 30		
ρ	**LRT**	F_{max}	F_{new}	**LRT**	F_{max}	F_{new}
0.1	0.0400	0.0244	0.0450	0.0228	0.0360	0.0436
0.2	0.0296	0.0228	0.0466	0.0328	0.0392	0.0502
0.3	0.0322	0.0264	0.0524	0.0326	0.0320	0.0456
0.4	0.0280	0.0286	0.0486	0.0330	0.0368	0.0462
0.5	0.0300	0.0306	0.0500	0.0242	0.0332	0.0434
0.6	0.0250	0.0266	0.0464	0.0250	0.0346	0.0474
0.7	0.0228	0.0296	0.0522	0.0210	0.0368	0.0468
0.8	0.0160	0.0252	0.0472	0.0142	0.0362	0.0442
0.9	0.0166	0.0272	0.0510	0.0128	0.0358	0.0424

Table 2. Rejection Proportions (α = 0.05)

			k = 5			k = 30		
ρ_1	ρ_2	ρ_3	LRT	F_{max}	F_{new}	LRT	F_{max}	F_{new}
0.1	0.3	0.6	0.1290	0.1420	0.2312	0.9182	0.9284	0.9396
0.1	0.3	0.8	0.4310	0.5832	0.6954	0.9736	1.0000	1.0000
0.1	0.5	0.4	0.0794	0.0810	0.1542	0.7454	0.6638	0.7186
0.1	0.5	0.6	0.1472	0.1452	0.2654	0.9316	0.9294	0.9450
0.1	0.5	0.8	0.4134	0.5306	0.6652	0.9764	1.0000	1.0000
0.1	0.6	0.2	0.1392	0.1644	0.2620	0.9318	0.9512	0.9614
0.1	0.6	0.4	0.1354	0.1370	0.2402	0.9164	0.9180	0.9318
0.1	0.6	0.6	0.1752	0.1882	0.3288	0.9604	0.9698	0.9832
0.1	0.6	0.8	0.4122	0.5172	0.6700	0.9780	1.0000	1.0000
0.1	0.7	0.2	0.2500	0.3150	0.4448	0.9682	0.9992	0.9994
0.1	0.7	0.4	0.2386	0.2756	0.4108	0.9690	0.9968	0.9962
0.1	0.7	0.6	0.2518	0.2806	0.4578	0.9716	0.9968	0.9982
0.1	0.7	0.8	0.4558	0.5534	0.7364	0.9774	1.0000	1.0000
0.3	0.3	0.2	0.0290	0.0252	0.0534	0.0846	0.0684	0.0902
0.3	0.3	0.4	0.0336	0.0254	0.0530	0.0904	0.0874	0.1048
0.3	0.3	0.6	0.0896	0.0860	0.1458	0.7282	0.7184	0.7524
0.3	0.3	0.8	0.3786	0.4610	0.5968	0.9988	1.0000	1.0000
0.3	0.5	0.4	0.0428	0.0380	0.0756	0.2446	0.2254	0.2708
0.3	0.5	0.6	0.0818	0.0768	0.1432	0.6008	0.6058	0.6564
0.3	0.5	0.8	0.3206	0.3696	0.4984	0.9978	0.9994	0.9998
0.3	0.6	0.2	0.1052	0.1076	0.1842	0.8546	0.8410	0.8650
0.3	0.6	0.4	0.0770	0.0642	0.1276	0.6194	0.6144	0.6488
0.3	0.6	0.6	0.0968	0.0910	0.1706	0.7276	0.7232	0.7886
0.3	0.6	0.8	0.3090	0.3502	0.4932	0.9974	0.9996	0.9998
0.3	0.7	0.2	0.2158	0.2444	0.3632	0.9866	0.9942	0.9954
0.3	0.7	0.4	0.1566	0.1628	0.2594	0.9426	0.9596	0.9688
0.3	0.7	0.6	0.1554	0.1592	0.2730	0.9226	0.9376	0.9526
0.3	0.7	0.8	0.3392	0.3662	0.5462	0.9984	1.0000	1.0000
0.5	0.3	0.4	0.0424	0.0384	0.0704	0.2348	0.2226	0.2598
0.5	0.3	0.6	0.0774	0.0756	0.1346	0.5986	0.5948	0.6466
0.5	0.3	0.8	0.3306	0.3816	0.5086	0.9986	0.9996	1.0000
0.5	0.5	0.4	0.0350	0.0316	0.0570	0.0936	0.0998	0.1276
0.5	0.5	0.6	0.0340	0.0334	0.0610	0.1144	0.1418	0.1690
0.5	0.5	0.8	0.2116	0.2464	0.3606	0.9730	0.9916	0.9924
0.5	0.6	0.2	0.1132	0.1152	0.1998	0.8388	0.8172	0.8516
0.5	0.6	0.4	0.0478	0.0422	0.0860	0.3062	0.3314	0.3778
0.5	0.6	0.6	0.0386	0.0342	0.0684	0.1150	0.1402	0.1756
0.5	0.6	0.8	0.1694	0.1986	0.2982	0.9350	0.9782	0.9828
0.5	0.7	0.2	0.1996	0.2070	0.3240	0.9750	0.9838	0.9854
0.5	0.7	0.4	0.1068	0.1054	0.1812	0.7638	0.8224	0.8454

Table 2. Continued

			k = 5			k = 30		
ρ_1	ρ_2	ρ_3	LRT	F_{max}	F_{new}	LRT	F_{max}	F_{new}
0.5	0.7	0.6	0.0580	0.0574	0.1142	0.4240	0.5082	0.5506
0.5	0.7	0.8	0.1648	0.1808	0.3036	0.9100	0.9668	0.9718
0.7	0.3	0.2	0.1966	0.2346	0.3370	0.9868	0.9940	0.9966
0.7	0.3	0.4	0.1582	0.1680	0.2686	0.9534	0.9690	0.9746
0.7	0.3	0.6	0.1536	0.1550	0.2734	0.9302	0.9452	0.9604
0.7	0.3	0.8	0.3616	0.3926	0.5638	0.9980	1.0000	1.0000
0.7	0.5	0.2	0.1924	0.2068	0.3222	0.9770	0.9856	0.9878
0.7	0.5	0.4	0.1104	0.1088	0.1890	0.7676	0.8230	0.8476
0.7	0.5	0.6	0.0594	0.0644	0.1152	0.4272	0.5164	0.5638
0.7	0.5	0.8	0.1578	0.1838	0.3014	0.9238	0.9696	0.9750
0.7	0.6	0.2	0.2146	0.2156	0.3686	0.9800	0.9860	0.9888
0.7	0.6	0.4	0.1088	0.1116	0.1944	0.7580	0.8078	0.8348
0.7	0.6	0.6	0.0396	0.0398	0.0732	0.1454	0.2096	0.2416
0.7	0.6	0.8	0.0852	0.0976	0.1700	0.6738	0.8080	0.8364
0.7	0.7	0.2	0.2650	0.2754	0.4550	0.9914	0.9952	0.9980
0.7	0.7	0.4	0.1354	0.1346	0.2554	0.8650	0.8924	0.9240
0.7	0.7	0.6	0.0358	0.0394	0.0742	0.1648	0.2256	0.2772
0.7	0.7	0.8	0.0386	0.0520	0.0950	0.2916	0.4322	0.4732
0.9	0.3	0.2	0.6918	0.9104	0.9448	0.9978	1.0000	1.0000
0.9	0.3	0.4	0.6914	0.8582	0.9058	0.9996	1.0000	1.0000
0.9	0.3	0.6	0.6584	0.8110	0.8774	0.9996	1.0000	1.0000
0.9	0.3	0.8	0.6780	0.7952	0.8918	0.9998	1.0000	1.0000
0.9	0.5	0.2	0.6888	0.8716	0.9154	0.9962	1.0000	1.0000
0.9	0.5	0.4	0.6394	0.7738	0.8564	1.0000	1.0000	1.0000
0.9	0.5	0.6	0.5408	0.6740	0.7706	1.0000	1.0000	1.0000
0.9	0.5	0.8	0.5152	0.6016	0.7428	0.9998	1.0000	1.0000
0.9	0.6	0.2	0.6806	0.8504	0.9114	0.9972	1.0000	1.0000
0.9	0.6	0.4	0.6200	0.7356	0.8200	1.0000	1.0000	1.0000
0.9	0.6	0.6	0.4552	0.5694	0.6916	0.9996	1.0000	1.0000
0.9	0.6	0.8	0.3686	0.4440	0.5958	0.9964	0.9998	0.9998
0.9	0.7	0.2	0.6782	0.8318	0.8934	0.9974	1.0000	1.0000
0.9	0.7	0.4	0.5848	0.7014	0.8020	1.0000	1.0000	1.0000
0.9	0.7	0.6	0.3790	0.4932	0.6234	0.9980	1.0000	1.0000
0.9	0.7	0.8	0.2014	0.2706	0.3984	0.9666	0.9970	0.9974

Example with Real Life Data

In this section, three tests using real life data collected from Srivastava and Katapa (1986) are prepared. The data is split randomly into three samples. Table 3 gives the values of pattern intensity on soles of feet in fourteen families, where values for daughters and sons are combined.

Table 3. Values of pattern intensity on soles of feet in 14 families

	Family #	Mother	Father	# Siblings	Siblings Values
	2	2	3	2	2, 3
	7	4	3	7	2, 2, 3, 6, 3, 5, 4
Sample A	8	3	7	7	2, 4, 7, 4, 4, 7, 8
	11	5	6	4	5, 3, 4, 4
	14	2	3	3	2, 2, 2
	1	2	3	2	2, 2
	5	2	3	2	6, 6
Sample B	6	4	3	3	4, 3, 3
	9	5	5	2	5, 6
	13	6	3	4	4, 3, 3, 3
	3	2	3	3	2, 2, 2
Sample C	4	2	4	5	2, 2, 2, 2, 2
	10	5	4	3	4, 5, 4
	12	2	4	2	2, 4

First ignoring the father's and mother's values, transform the siblings' values by multiplying each observation vector by Helmert's orthogonal matrix Q

where where $$\underset{p_i x p_i}{Q} = \begin{bmatrix} \frac{1}{\sqrt{p_i}} & \frac{1}{\sqrt{p_i}} & \frac{1}{\sqrt{p_i}} & \cdots & \frac{1}{\sqrt{p_i}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{p_i(p_i-1)}} & \frac{1}{\sqrt{p_i(p_i-1)}} & \frac{1}{\sqrt{p_i(p_i-1)}} & \cdots & \frac{(p_i-1)}{\sqrt{p_i(p_i-1)}} \end{bmatrix}$$

This gives transformed vectors $\underset{\sim}{u_i}, \underset{\sim}{v_j}$ *and* $\underset{\sim}{w_l}$ respectively for $i = 1, 2, ..., k_1$; $j = 1, 2, ..., k_2$ and $l = 1, 2, ..., k_3$. Here, $k_1 = 5, k_2 = 5$ and $k_3 = 4$.
Srivastava's formula, given by (3), is used to compute intraclass correlation coefficients. The computed values of intraclass correlation coefficients are $\hat{\rho}_1 = 0.5895$, $\hat{\rho}_2 = 0.9159$ and $\hat{\rho}_3 = 0.7685$ and $\hat{\rho}$ =0.4923.

Table 4. Raw Computations

				i	***i**α/6**	***P*(*i*)**	**Col3 < Col2?**
F1	9.2874	P1	0.000014	1	0.008333	0.000014	yes
F2	0.1355	P5	0.003139	2	0.016667	0.003139	yes
F3	0.1640	P6	0.003859	3	0.025000	0.003859	yes
F4	0.1077	P3	0.996140	4	0.033333	0.996140	no
F5	7.3798	P2	0.996860	5	0.041667	0.996860	no
F6	6.0994	P4	0.999990	6	0.050000	0.999990	no

Note. Conclusion = Reject.

The computed values of LRT statistic and $F_{\max}$ statistic obtained from formula (8) and (9) respectively are as follows:

Table 5. Test Statistics and their Critical Values

	Test Statistic	**CV (α = 0.01)**	**CV (α = 0.05)**	**CV (α = 0.10)**
LRT	7.7820	9.2103	5.9915	4.6052
$F_{\max}$	10.4510	10.1660	6.2630	5.0025
F_{new}	N/A	Reject	Reject	Reject

References

Bhandary, M. and Alam, M. K. (2000). Test for the equality of intraclass correlation coefficients under unequal family sizes. *Communications in Statistics – Theory and Methods, 29*(4): 755-768.

Bhandary, M. and Fujiwara, K. (2006). A small sample test for the equality of intraclass correlation coefficients under unequal family sizes for several populations. *Communications in Statistics – Simulation and Computation, 35*(3): 765-778.

Donner, A. and Bull, S. (1983). Inferences concerning a common intraclass correlation coefficient. *Biometrics, 39*: 771-775.

Donner, A. and Zou, G. (2002). Testing the equality of dependent intraclass correlation coefficients. *The Statistician*, *51*(3): 367-379.

Gokhale, D.V. and SenGupta, A. (1986). Optimal tests for the correlation coefficient in a symmetric multivariate normal population. *J. Statist. Plann. Inference*, *14*: 263-268.

Huang, W. and Sinha, B. K. (1993). On optimum invariant tests of equality of intraclass correlation coefficients. *Annals of the Institute of Statistical Mathematics*, *45*(3): 579-597.

Konishi, S. (1985). Normalizing and variance stabilizing transformations for intraclass correlations. *Annals of the Institute of Statistical Mathematics*, *37*: 87-94.

Konishi, S. and Gupta, A. K. (1989). Testing the equality of several intraclass correlation coefficients. *J. Statist. Plann. Inference*, *21*: 93-105.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. New York: Wiley.

Rosner, B., Donner, A. and Hennekens, C. H. (1977). Estimation of intraclass correlation from familial data. *Applied Statistics*, *26*: 179-187.

Rosner, B., Donner, A. and Hennekens, C. H. (1979). Significance testing of interclass correlations from familial data. *Biometrics*, *35*: 461-471.

SenGupta, A. (1988). On loss of power under additional information – an example. *Scand. J. Statist*., *15*: 25-31.

Scheffe, H. (1959). *The Analysis of Variance.* New York: Wiley.

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. *Canadian Journal of Statistics*, *14*: 29-42.

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

Velu, R. and Rao, M. B. (1990). Estimation of parent-offspring correlation. *Biometrika*, *77*(3): 557-562.

Young, D. and Bhandary, M. (1998). Test for the equality of intraclass correlation coefficients under unequal family sizes. *Biometrics*, *54*(4): 1363-1373.