66

Chinese Journal of Psychology
1991, Vol. 33, 3-10

中華心理學刊
民 80, 33 卷, 3-10

ON NON-NULL TESTS OF INTRACLASS CORRELATIONS

PETER H. SCHÖNEMANN
*National Taiwan University**

It is shown that the *non-null* distribution of a simple transform of the 2-variable intraclass correlation is *central F*. Hence, in principle, standard F-tables could be used to test hypotheses about intraclass correlations. In a simulation, Fisher's normal approximation was compared with the exact test for three different intraclass correlation estimates. Overall, the results were very close, confirming that Fisher's z-transform is an excellent approximation even when the sample size is small.

Intraclass correlations are widely used in human genetics to assess the similarity of various types of twins and siblings. In these applications, N, the number of families, is usually quite small (e. g., Shields, 1962; Jinks and Fulker, 1970). Early authors, e. g., Harris (1913), Fisher (1921), had defined the intraclass correlation r_I' as the product-moment correlation of a "double entry table", a $2N \times 2$ table of scores where each pair is entered once as (x_i, y_i) and once as (y_i, x_i), to remove the ambiguity whether a twin's score should occur on the right or on the left. Harris (1913) has shown that this definition is equivalent to

$$r_I' := \frac{(SSB - SSW)}{(SSB + SSW)} = r' \qquad (1.1)$$

where SSB, SSW are the between and within sums of squares of the original $N \times 2$ score matrix. (The notation r_I' follow Snedecor and Cochran, 1967. For notational convenience, it will be denoted r' from now on. The generalization to classes consisting of $m > 2$ siblings will not be considered here).

Some modern authors (e. g., Winer, 1971, p. 244; Jinks and Fulker, 1970) follow Scheffe (1959, p. 233) in restricting the definition of the intraclass correlation to a variance ratio, which, of course, cannot be negative. In contrast, the intraclass correlation (1.1) will be negative whenever $SSB < SSW$. As Snedecor and Cochran (1967, p. 294) observed, "the [intraclass correlation] model is more general than the components of variance model... if, for instance, four young animals in a pen compete for an insufficient supply of food, the stronger animals may drive away the weaker and may regularly get most of the food. For this reason the variance in weight within pens may be larger than that between pens, this being a real phenomenon and not an accident of sampling. We say that there is a negative

* On Sabbatical leave from Purdue University. I would like to thank Professors C. M. Cheng and J. T. Huang for arranging this Sabbatical for me, the National Science Council for funding it, and the students and faculty of the NTU Psychology Department for their hospitality, which made it so enjoyable. Requests for reprints should be addressed to Peter H. Schönemann, Department of Psychological Sciences, Purdue University, West Lafayette IN 47906, USA.

correlation between the weights within a pen." The analogy to twins is obvious. Hence it should not come as a surprise that human geneticists occasionally encounter negative intraclass correlations (e.g., Jinks and Fulker, loc. cit., p. 333).

In (1921), Fisher derived the exact distribution of r'.

$$dF(r')=c(N,\pi)\cdot\operatorname{sech}^{N-.5}(z-\theta)\cdot\exp[-.5(z-\theta)] \quad (1.2)$$

where

$$c(N,\pi),=\Gamma(N-.5)/\Gamma(N-1)\sqrt{2\pi};$$
$$z:=\tanh^{-1}(r'),$$
$$\theta:=\tanh^{-1}(\rho),$$

r' is the sample intraclass correlation, ρ its population value, and N the number of classes (Kendall and Stewart, 1961, p. 315). Since this distribution is not easily tabled, Fisher recommended his well-known z-transformation, with $\sigma_r{}'=(N-3/2)^{-1/2}$, as an approximate test for the non-null case when N is "large."

Geneticists often use a variant of $r_I{}'$ based on mean squares,

$$r_I:=\frac{(MSB-MSW)}{(MSB+MSW)}=r \quad (1.3)$$

which relates directly to the F of a random effects ANOVA. As Snedecor and Cochran note, "$r_I{}'$ differs slightly from r_I, the difference being trivial unless [N] (the number of classes) is small" (p. 296, loc. cit.). The exact relation is

$$r'=(r-a)/(1-ar)<r<r'+a,$$

where

$$a:=1/(2N-1). \quad (1.4)$$

Although r has no longer an intuitively straightforward interpretation as the corrrelation of a double entry table, it could be viewed simply as an approximation to r'.

In some applications, it may not be unreasonable to assume that the population mean μ is "known", for example, when the sample size is large, or when standard tests are used whose mean is known. In twin research the Wechsler, the Stanford Binet, and the PMA are often used whose population means are 100. To cover these situations, a third intraclass correlation,

$$r_I{}^*:=\frac{(SSB^*-SSW)}{(SSB^*+SSW)}=r^*, \quad (1.5)$$

will be defined, where SSB^* is based on the deviation scores around the known population mean μ. SSW is the same as in (1.2) since it is unaffected by the mean.

The well-known definitions of the sums of squares reduce for $N\times2$ tables to

$$SSW:=(x-y)'(x-y)/2 \quad (1.6)$$

where x, y denote the two columns of the matrix of deviation score vectors around the (sample) grand mean (or, equivalently, around the population mean, since the mean cancels in difference scores). Similarly, the between sums of squares are given by

$$SSB:=(x+y)'(x+y)/2 \quad (1.7)$$

when the mean is not known, so that x, y are the deviation score vectors around the (sample) grand mean. They are given by

$$SSB^*:=(x^*+y^*)'(x^*+y^*)/2 \quad (1.8)$$

when the mean is known, so that x^*, y^* are the deviation score vectors around the population mean μ.

It will now be shown that the non-null distributions of a simple transform of the three intraclass correlations, (1.2), (1.3), (1.5) are central F. The transfrom is given by the involution

$$s:\ Re \to Re:\ s(x):=(1-x)/(1+x),\quad x \neq -1, \tag{1.9}$$

which, as is easily verified, satisfies

$$s[s(x)]=x,\ s(x)+s(1/x)=0,\quad s(x)s(-x)=1. \tag{1.10}$$

This involution reduces the three intraclass correlations to

$$r'=s(SSW/SSB),\quad r=s(MSW/MSB),\quad r^*=s(SSW/SSB^*). \tag{1.11}$$

In view of (1.10), one finds

$$s(r')=SSW/SSB,\quad s(r)=MSW/MSB,\quad s(r^*)=SSW/SSB^*. \tag{1.12}$$

Thus, even though, strictly speaking, it is incorrect to interpret intraclass correlations as variance ratios, the s-transforms of r and r^* can be so interpreted, and the s-transform of r' is proportional to a variance ratio.

EXACT NON-NULL DISTRIBUTIONS

The basic argument needed to establish that the s-transforms of the intraclass correlations (1.12) are central F follows along standard lines: A random vector (x_i, y_i) is transformed to principal axes and then standardized. The resulting vector (u_i, v_i) is $N(\phi, I)$. Hence any pair of quadratic forms (sums of squares) in the u_i and v_i, respectively, are two independent chi-square variables and their ratio is central F.

Theorem 1: Let $(x_i, y_j)'$ be N independent, identically distributed random vectors with bivariate normal distribution

$$N(\mu;\ \Sigma),$$

where

$$\mu':\ =(\mu,\ \mu)$$

and

$$\Sigma:\ \sigma^2\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}. \tag{2.1}$$

Then

$$F(r'):\ =Ns(\rho)/s(r')(N-1) =s(\rho)MSB/MSW \sim F_{N-1,N}, \tag{2.2}$$

$$F(r):\ =s(\rho)/s(r)=F(r') =s(\rho)MSB/MSW \sim F_{N-1,N}, \tag{2.3}$$

$$F(r^*):\ =s(\rho)/s(r^*) =s(\rho)SSB^*/SSW \sim F_{N,N}, \tag{2.4}$$

where $F_{p,q}$ denotes a central F-variate with p degrees of freedom (df) for the numerator and q df for the denominator.

Proof: If the eigendecomposition of $\Sigma=\sigma^2 VD^2V'$, where

$$V:\ =2^{-1/2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix},\quad D:\ =\begin{pmatrix}(1+\rho)^{1/2} & 0\\ 0 & (1-\rho)^{1/2}\end{pmatrix}, \tag{2.5}$$

then the principal axes transformation, followed by standardization,

$$D^{-1}V\begin{pmatrix}x_i-\mu\\ y_i-\mu\end{pmatrix}/\sigma =\begin{pmatrix}u_i\\ v_i\end{pmatrix}\sim N(\phi,\ I) \tag{2.6}$$

produces two uncorrelated standard normal variables u_i, v_i, which reproduce the original variables as

$$x_i=\sigma[(1+\rho)/2]^{1/2}u_i + \sigma[(1-\rho)/2]^{1/2}v_i+\mu$$
$$y_i=\sigma[(1+\rho)/2]^{1/2}u_i - \sigma[(1-\rho)/2]^{1/2}v_i+\mu. \quad (2.7)$$

Hence

$$x_i-y_i=\sigma[2(1-\rho)]^{1/2}v_i \quad (2.8)$$

$$x_i+y_i=\sigma[2(1+\rho)]^{1/2}u_i+2\mu, \quad (2.9)$$

and the sum of the two sample means (= twice the grandmean) is

$$\bar{x}+\bar{y}=\sigma[2(1+\rho)]^{1/2}\bar{u}+2\mu. \quad (2.10)$$

Therefore, the sum of deviation scores around the grandmean is

$$(x_i-\bar{x})+(y_i-\bar{y}) =\sigma[2(1+\rho)]^{1/2}(u_i-\bar{u}). \quad (2.11)$$

On summing the squares of (2.7), (2.12), and of (2.8) when $\mu=0$, one obtains the central chi-square variates,

$$2SSW=\Sigma_i(x_i-y_i)^2 \sim 2\sigma^2(1-\rho)\chi^2_N,$$
$$2SSB^*=\Sigma_i(x_i+y_i)^2 \sim 2\sigma^2(1+\rho)\chi^2_N$$
$$2SSB=\Sigma_i(x_i+y_i-\bar{x}-\bar{y})^2 \sim 2\sigma^2(1+\rho)\chi^2_{N-1}, \quad (2.12)$$

where χ_n^2 denotes a central chi-square variate with n *df*. Since SSW and SSB^* are quadratic forms of rank N, they each have N *df*. Since SSB is a quadratic form of rank $N-1$, it has $N-1$ *df*. Since SSW is a function of the v_i, while SSB, SSB^* are functions of the u_i which are independent of the v_i, SSB and SSB^* are independent of SSW. Hence the ratios (2.2), (2.3) (2. 4) are central F with the indicated *df*. qed.

SIMULATIONS

The performance of the approximate Fisher z-tests of r^*, r', and r was compared with that of the s-transforms in an (*IBM PC*) computer simulation. As sample sizes N=20, 40, 80, and 160 were used, and as population correlations ρ= .3, .5, .7, and .9. Although the means and variances should of course not affect the outcome of the tests, two different means, $\mu=5$ and $\mu=-20$, and two standard deviations, $\sigma=3$ and $\sigma=15$, were used as a check on the robustness of the simulation results. A completely crossed design for N, ρ, μ, and σ thus yielded (4)(4)(2)(2)=64 parameter configurations, each of which was replicated 400 times.

Within each replication, an $N\times2$ score matrix was drawn randomly from a normal (0, 1) distribution and then transformed according to eq. (2.7) into a score matrix of N replications with specified population mean, variance and intraclass correlation. After computing the intraclass correlations (1.1), (1.3), (1.5), the rejection rates were tabulated for five test-statistics: the two exact tests $F(r')$ $(=F(r))$ and $F(r^*)$, and the three approximate tests $z(r')$, $z(r)$, and $z(r^*)$, based on Fisher's inverse hyperbolic tangent transform of r', r, and r^*, respectively, with $\sigma_r:=(N-1.5)^{-1/2}$.

The results of this simulation are summarized in Table 1 for four different α levels (.2, .1, .05, and .025). The first five columns in the body of each subtable give the empirical rejection rates and the last five columns the differences, observed minus nominal α, for ease of comparison. The next to last row in each subtable gives the column means (across correlations nested within sample sizes) as an overall index of bias. The last row gives the column sums of squars of the observed minus nominal α discrepancies as an index of overall performance.

Each entry in the body of the subtables is based on 1600 replications (it represents the average across four means by variances configurations), Inspite of this relatively large replication number, there is still some variability left in the observed rejection rates but most of it appears to be unrelated to the systematically varied parameters N, ρ, μ, and σ. In particular, the rejection rates of

Table 1
Rejection rates of various tests of intraclass correlations

α=.20		Observed rates					Observed minus nominal rates				
N	ρ	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$
20	.3	.198	.196	.193	.164	.194	−.002	−.004	−.007	−.036	−.007
20	.5	.209	.203	.203	.181	.200	.009	.002	.002	−.020	.000
20	.7	.219	.222	.211	.185	.217	.018	.022	.010	−.015	.017
20	.9	.190	.197	.187	.163	.195	−.010	−.003	−.013	−.038	−.005
40	.3	.202	.204	.200	.179	.201	.002	.004	.000	−.021	.001
40	.5	.199	.196	.198	.175	.194	−.001	−.004	−.002	−.025	−.006
40	.7	.200	.197	.194	.169	.194	.000	−.003	−.006	−.031	−.006
40	.9	.198	.202	.195	.177	.201	−.002	.002	−.005	−.023	.001
80	.3	.192	.194	.192	.182	.193	−.008	−.006	−.008	−.018	−.007
80	.5	.220	.220	.219	.199	.219	.020	.020	.018	−.002	.019
80	.7	.201	.197	.200	.181	.196	.001	−.003	.000	−.019	−.004
80	.9	.194	.198	.192	.178	.198	−.006	−.003	−.009	−.022	−.003
160	.3	.220	.218	.220	.209	.218	.020	.018	.020	.010	.018
160	.5	.191	.187	.189	.177	.187	−.009	−.013	−.011	−.024	−.013
160	.7	.221	.220	.219	.206	.219	.020	.020	.019	.005	.019
160	.9	.200	.200	.200	.188	.200	−.001	.000	−.001	−.013	.000
Means		.203	.203	.201	.182	.202	.003	.003	.000	−.018	.001
SSQ							.002	.002	.002	.008	.002

α=.10		Observed rates					Observed minus nominal rates				
N	ρ	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$
20	.3	.087	.081	.084	.061	.078	−.013	−.020	−.016	−.039	−.022
20	.5	.097	.100	.093	.075	.097	−.004	.001	−.008	−.025	−.003
20	.7	.109	.111	.106	.088	.109	.009	.011	.005	−.013	.009
20	.9	.100	.094	.092	.074	.093	−.001	−.005	−.008	−.027	−.007
40	.3	.103	.101	.100	.083	.101	.002	.001	.001	−.017	.001
40	.5	.104	.100	.102	.086	.099	.004	−.001	.002	−.014	−.001
40	.7	.091	.089	.090	.074	.087	−.009	−.011	−.010	−.026	−.013
40	.9	.098	.095	.096	.083	.095	−.002	−.004	−.004	−.017	−.004
80	.3	.105	.104	.105	.100	.104	.005	.004	.005	.001	.004
80	.5	.114	.115	.114	.110	.115	.014	.015	.014	.010	.014
80	.7	.098	.098	.098	.090	.098	−.002	−.002	−.002	−.010	−.002
80	.9	.091	.093	.089	.084	.092	−.009	−.007	−.011	−.017	−.008
160	.3	.109	.108	.109	.104	.108	.009	.008	.009	.004	.008
160	.5	.097	.099	.097	.091	.098	−.003	−.002	−.003	−.009	−.002
160	.7	.112	.112	.112	.104	.112	.012	.012	.012	.004	.012
160	.9	.085	.086	.084	.081	.086	−.015	−.014	−.016	−.019	−.015
Means		.100	.099	.098	.087	.098	.000	−.001	−.002	−.013	−.002
SSQ							.001	.001	.001	.005	.002

(Table 1 *continued*)

$\alpha=.05$		Observed rates					Observed minus nominal rates				
N	ρ	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$
20	.3	.036	.039	.035	.030	.038	−.014	−.011	−.015	−.021	−.013
20	.5	.044	.043	.042	.030	.043	−.006	−.007	−.008	−.021	−.008
20	.7	.056	.057	.053	.044	.056	.005	.007	.003	−.007	.006
20	.9	.046	.047	.045	.038	.044	−.004	−.003	−.005	−.012	−.006
40	.3	.047	.047	.047	.041	.046	−.004	−.003	−.004	−.009	−.004
40	.5	.052	.053	.051	.045	.052	.002	.002	.001	−.005	.002
40	.7	.044	.044	.044	.041	.044	−.096	−.006	−.006	−.009	−.006
40	.9	.053	.055	.053	.049	.055	.003	.005	.003	−.001	.005
80	.3	.053	.053	.053	.050	.053	.003	.003	.003	.000	.003
80	.5	.058	.055	.057	.050	.055	.008	.004	.007	.000	.004
80	.7	.051	.050	.051	.046	.049	.001	.000	.001	−.005	−.001
80	.9	.047	.045	.047	.045	.045	−.002	−.004	−.002	−.006	−.004
160	.3	.057	.059	.057	.053	.058	.007	.009	.007	.003	.008
160	.5	.047	.048	.047	.045	.048	−.003	−.002	−.003	−.005	−.002
160	.7	.054	.054	.054	.050	.054	.004	.004	.004	.000	.004
160	.9	.042	.043	.040	.039	.043	−.008	−.007	−.010	−.011	−.007
Means		.049	.049	.049	.043	.049	−.001	−.001	−.001	−.007	−.001
SSQ							.001	.000	.001	.001	.001

$\alpha=.025$		Observed rates					Observed minus nominal rates				
N	ρ	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$	$F(r^*)$	$F(r')$	$z(r^*)$	$z(r')$	$z(r)$
20	.3	.019	.019	.017	.016	.019	−.007	−.007	−.008	−.009	−.007
20	.5	.025	.022	.022	.016	.022	.000	−.003	−.002	−.009	−.003
20	.7	.027	.029	.025	.018	.029	.002	.004	.000	−.007	.004
20	.9	.025	.027	.024	.021	.027	.000	.002	−.001	−.004	.002
40	.3	.025	.022	.024	.020	.022	.000	−.003	−.001	−.005	−.003
40	.5	.027	.028	.027	.024	.028	.002	.003	.002	−.001	.003
40	.7	.023	.022	.023	.021	.022	−.002	−.003	−.002	−.004	−.003
40	.9	.027	.031	.027	.025	.031	.002	.005	.002	−.001	.005
80	.3	.033	.030	.032	.026	.030	.008	.005	.007	.001	.005
80	.5	.032	.031	.031	.025	.031	.007	.005	.006	.000	.005
80	.7	.022	.023	.022	.020	.023	−.003	−.002	−.004	−.005	−.002
80	.9	.029	.028	.029	.027	.028	.004	.003	.004	.002	.003
160	.3	.030	.029	.030	.027	.029	.005	.004	.005	.002	.004
160	.5	.026	.026	.026	.024	.026	.001	.001	.001	−.001	.001
160	.7	.024	.023	.024	.023	.023	−.001	−.002	−.001	−.002	−.002
160	.9	.022	.022	.022	.019	0.22	−.003	−.003	−.003	−.006	−.003
Means		.026	.026	.025	.022	.026	.001	.001	.000	−.003	.001
SSQ							.000	.000	.000	.000	.000

Fisher's z-test are virtually indistinguishable from those of the exact F-tests with one exception, those of ($z(r')$, which, for small N, tends to accept H_0 too readily.

Thus, the overall conclusion is that Fisher's z-tests based on r^* or r are excellent approximations to the corresponding exact F-tests even when sample sizes are small (as they often are in twin studies). For small sample sizes, the intraclass correlation r based on mean squares is preferable over the more descriptive estimate r' based on sums of squares, because the z-transform of r' has a slight conservative bias when $N<40$.

NUMERICAL ILLUSTRATION

To test

$$H_0\text{: } \rho=\rho_0 \ vs.\ H_1\text{: } \rho>\rho_0, \qquad (3.2)$$

in the most typical case, when μ is not known, first compute the sample intraclass correlation r

$$r=\frac{(MSB-MSW)}{(MSB+MSW)} \qquad (3.2)$$

Then the statistic

$$F(r)=\frac{s(\rho_0)}{s(r)}=\frac{(1-\rho_0)(1+r)}{(1-r)(1+\rho_0)} \qquad (3.3)$$

is central F with $N-1$ df for the numerator and N df for the denominator under H_0. If this statistic exceeds the tabled F-value at the alpha level, reject H_0

If x_i, y_i now denote the observed scores, then the sums of squares are given by

$$SSB\text{: } =\sum_i[(x_i-\bar{w})+(y_i-\bar{w})]^2/2,$$
$$SSW\text{:} =\sum_i(x_i-y_i)^2/2 \qquad (3.4)$$

where $\bar{w}$ denotes the grandmean, $\bar{w}$: $=(\sum_i x_i+\sum_i y_i)/2N$. To illustrate these computations concretely, assume the mean is not known and $r=.8$, $N=20$, $\rho_0=.6$. Let us use $\alpha=.05$. Then $s(r)=(1-.8)/(1+.8)=.111$, $s(\rho_0)=.25$, and $F(r)=.25/.11=2.27$. Since the tabled .05 upper tail value of F with (19, 20) degrees of freedom is 2.12, H_0: $\rho=.6$ can be rejected in favor of H_1: $\rho>.6$ at the .05 level of significance.

A $100(1-\alpha)\%$ confidence interval is obtained in the usual way by inverting the probability over the acceptance region,

$$\mathrm{Prob}(L_1<F(r)<L_2)=1-\alpha, \qquad (3.5)$$

where L_2: $=F^{-1}{}_{N-1,N}(1-\alpha/2)$ and L_1: $=F^{-1}{}_{N-1,N}(\alpha/2)$. Since

$$L_1<s(\rho)/s(r)<L_2 \Leftrightarrow L_1 s(r)<s(\rho)<L_2 s(r), \qquad (3.6)$$

one obtains, upon applying the involution s,

$$\mathrm{Prob}\{s[s(r)L_2]<\rho<s[s(r)L_1]\} =1-\alpha, \qquad (3.7)$$

e.g., to continue with the above example, where $F^{-1}{}_{19,20}=.464$,

$$\mathrm{Prob}\{s[(.111)(2.12)]<\rho <s[(.111)(.464)]\} =\mathrm{Prob}[s(.235)<\rho<s(.052)] =\mathrm{Prob}[.62<\rho<.90]=.90. \qquad (3.8)$$

This 90% confidence interval agrees with that obtained via Fisher's z-transform within two decimal places.

REFERNENCES

FISHER, R. A. (1921). On the "probable error" of a correlation deduced from a small sample. *Metron*, **1**, 3-32.

HARRIS, J. A. (1913). On the calculation of intra-class and inter-class coefficients of correlation from class moments when the number of possible combinations is large. *Biometrika*, **9**, 446.

JINKS, J. L. and FULKER, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Review*, **73**, 311-349.

KENDALL, M. G. and STEWART, A. (1961). *Advanced theory of statistics*. Volume II. New York: Hafner.

SCHEFFE, H. A. (1959). *The analysis of variance*. New York: Wiley.

SHIELDS, J. (1962). *Monozygotic twins*. London: Oxford University Press.

SNEDECOR, G. W. and COCHRAN, W. G. (1967). *Statistical Methods*. Ames, Iowa: The Iowa State University Press, 6th edition.

WINER, B. J. (1971). *Statistical principles of experimental design*. Second edition. New York: McGraw Hill.

Received: *May 12, 1992*
Revised: *July 15, 1992*
Accepted: *July 18, 1992*

組內相關之非虛無檢驗

舒 彼 德
國立臺灣大學心理學系

本文顯示兩變項的組內相關經過簡單轉換後，其非虛無分配爲中央F分配。因此，一般通用的F表可以用來檢驗有關組內相關係數的假設。文中模擬研究比較費雪常態近似值的檢驗與三個不同組內相關係數估計值的精準檢驗，發現四者結果極爲相近。由此可知費雪的Z轉換是極佳的近似估計，卽使在小樣本的情況下亦然。