Ann. Inst. Statist. Math.
37 (1985), Part A, 87-94

NORMALIZING AND VARIANCE STABILIZING TRANSFORMATIONS FOR INTRACLASS CORRELATIONS

SADANORI KONISHI

(Received Feb. 8, 1984; revised May 31, 1984)

Summary

A concept of normalizing transformations of statistics is constructed on the basis of the rate of convergence to normality. The concept is applied to derive a normalizing transformation of a maximum likelihood estimate of intraclass correlation coefficient in a p-variate normal sample. Numerical comparisons are made to examine whether the proposed transformation is efficient to achieve normality. The relationship between normalization and variance stabilization is also considered.

1. Introduction

Fisher's z transformation [4] for the correlation coefficient, r, in a bivariate normal sample is of practical importance in determining significance points of the probability distribution or in constructing confidence intervals for population correlation coefficients. Transformations of this kind are, in general, made for two different purposes; partly for stabilizing an asymptotic variance, and partly for normalizing an asymptotic distribution. Fisher's z transformation for r is of particular interest, since the variance stabilization and normalization can be simultaneously achieved by the same transformation $z(r)=\frac{1}{2}\log\{(1+r)/(1-r)\}$.

An intraclass correlation coefficient is a measure of strength of relationship and useful in the estimation of the degree of resemblance between siblings in the biometrical study of inheritance. An explicit expression for the maximum likelihood estimate, r_i, of the intraclass correlation coefficient can be derived when the sample is drawn from a p-variate normal distribution and the number of siblings is the same in each family. It is known (Fisher [5], Chapter 7) that the variance

Key words and phrases: Fisher's z transformation, intraclass correlation, normalizing transformation, variance stabilization.

stabilizing transformation for r_i is for all values of p

$$(1.1) \qquad z_p(r_i)=\{(p-1)/(2p)\}^{1/2}\log[\{1+(p-1)r_i\}/(1-r_i)] .$$

The question is whether this transformation simultaneously yields a normalizing one. It must be remarked here that the concept of a normalizing transformation is ambiguous, while a variance stabilizing transformation is clear and succinct; we search for a function which renders an asymptotic variance independent of unknown parameters.

It is the purpose of this paper to construct a concept of normalizing transformations of statistics and to give a normalizing transformation for the intraclass correlation. The relationship between normalization and variance stabilization is also considered based on the transformation theory to be suggested here. Some numerical comparisons are made for approximate distributions of the intraclass correlation.

2. Normalization and variance stabilization

2.1. *Transformation theory*

Normalizing transformations of some statistics have been considered by Konishi [8], [9] and Efron [3]. We construct, on the basis of standpoints discussed in Bhattacharya and Ghosh [1] and Konishi [9], the concept of a normalizing transformation in the following way: Consider a random variable T_n whose distribution depends on the parameters n and $\theta=(\theta_1, \cdots, \theta_p)'$. Assume that there exist $\mu(\theta)$ and $\sigma(\theta)$ such that $\sqrt{n}\{T_n-\mu(\theta)\}/\sigma(\theta)$ has a limiting normal distribution with mean zero and variance 1 as n tends to infinity, and that the rate of convergence to normality is

$$\Pr[\sqrt{n}\{T_n-\mu(\theta)\}/\sigma(\theta)<x]=\Phi(x)+O(n^{-1/2}) ,$$

where $\Phi(x)$ is the standard normal distribution function. If there exists a strictly monotone function f such that

$$\Pr[\sqrt{n}\{f(T_n)-f(\mu(\theta))-c/n\}/\{\sigma(\theta)f'(\mu(\theta))\}<x]=\Phi(x)+O(n^{-1}) ,$$

where c is an asymptotic bias of the transformed variate $f(T_n)$, then $f(T_n)$ is said to be a normalizing transformation of T_n.

This means that by making a suitable transformation with an appropriate bias correction c, the term of $O(n^{-1/2})$ in an asymptotic expansion for the distribution of $f(T_n)$ can be made to vanish, so the error involved is of order n^{-1}. It is known (Rao [12], p. 385) that the problem of finding a variance stabilizing transformation is reduced to solving the differential equation $\sigma(\theta)f'(\mu(\theta))=1$ for a continuously differentiable function f in a neighborhood of $T_n=\mu(\theta)$.

Suppose that r is the sample correlation coefficient based on a sample of size $N=n+1$ from a bivariate normal distribution with correlation coefficient ρ. Taking $f(r)=z(r)=\frac{1}{2}\log\{(1+r)/(1-r)\}$ and $c=\rho/2$, we have (Konishi [9])

$$\Pr[\sqrt{n}\{z(r)-z(\rho)-\rho/(2n)\}<x]=\Phi(x)+O(n^{-1}),$$

while the remainder term of the limiting distribution of $\sqrt{n}(r-\rho)/(1-\rho^2)$ is $O(n^{-1/2})$. Hence the normalization and variance stabilization for r are simultaneously achieved by Fisher's z transformation.

Borges [2] has also used the same approach to obtain a normalizing transformation for a class of one-parameter distributions. The relationship between normalization and variance stabilization for a certain submodel of an n-dimensional exponential family has been considered by Hougaard [7].

2.2. *Transformations of intraclass correlations*

Let $X_\alpha=(x_{1\alpha}, \cdots, X_{p\alpha})'$, $\alpha=1, \cdots, N$, be a random sample of size $N=n+1$ from a p-variate normal distribution with mean vector μ and covariance matrix Σ. We assume that Σ has homogeneous variances and homogeneous covariances, so that

$$\Sigma=\sigma^2\{(1-\rho)I+\rho ee'\}$$

where $e=(1, \cdots, 1)'$ is the p-dimensional vector. For the mean vector μ we consider the following two cases; (i) elements of μ are all equal and (ii) μ is unrestricted. In the case where all of the elements of μ are equal, the maximum likelihood estimate, r_1, of the intraclass correlation coefficient ρ is

$$r_1=\sum_{\alpha=1}^{N}\sum_{i\neq j}^{p}(X_{i\alpha}-\bar{X})(X_{j\alpha}-\bar{X})\Big/\left\{(p-1)\sum_{\alpha=1}^{N}\sum_{i=1}^{p}(X_{i\alpha}-\bar{X})^2\right\}$$

where $\bar{X}=\sum_\alpha\sum_i X_{i\alpha}/(Np)$. Let $f(r_1)$ be a strictly monotone and twice continuously differentiable function in a neighborhood of $r_1=\rho$. Following a similar line of approach discussed in Fujikoshi [6] and Konishi [10] with the help of Theorem 2 of Bhattacharya and Ghosh [1], we obtain an asymptotic expansion for the distribution of $f(r_1)$ in the following;

$$(2.1)\quad \begin{aligned}&\Pr[\sqrt{N}\{f(r_1)-f(\rho)-c/N\}/\tau<x]\\&=\Phi(x)-N^{-1/2}[(7-5p)/\{3kp(p-1)\}-c/\tau+a(\rho, f', f'')x^2]\phi(x)\\&\quad+O(N^{-1})\end{aligned}$$

where $\phi(x)$ is the standard normal density function and

$$k=\sqrt{2}\{p(p-1)\}^{-1/2}, \qquad \tau^2=k^2(1-\rho)^2\{1+(p-1)\rho\}^2 f'(\rho)^2,$$

$$a(\rho, f', f'')=\left[\frac{2}{3}\{(p-2)-3(p-1)\rho\}+(1-\rho)\{1+(p-1)\rho\}f''(\rho)f'(\rho)^{-1}\right]\Big/ \{kp(p-1)\}.$$

We first search for a function which makes the coefficient $a(\rho, f', f'')$ of x^2 in (2.1) vanish. The problem of finding f is, noting that $f''(\rho)\cdot f'(\rho)^{-1}=d\{\log f'(\rho)\}/d\rho$, reduced to solving the differential equation

$$\frac{df}{d\rho}=(1-\rho)^{-2(2p-1)/(3p)}\{1+(p-1)\}\rho^{-2(1+p)/(3p)} \tag{2.2}$$

where $1>\rho>-(p-1)^{-1}$. The solution yields

$$z_2(r_1)=\frac{1}{2}\log\{(1+r_1)/(1-r_1)\} \qquad \text{for } p=2$$

and

$$B_p(r_1)=3(p-2)^{-1}[\{1+(p-1)r_1\}/(1-r_1)]^{(p-2)/(3p)} \qquad \text{for } p\geqq 3.$$

Calculating bias correction factors given by $c=\tau(7-5p)/\{3kp(p-1)\}$ in (2.1) for each transformation, we see that if $p=2$,

$$\Pr\left[\sqrt{N}\left\{z_2(r_1)-z_2(\rho)+\frac{1}{2}N^{-1}\right\}<x\right]=\Phi(x)+O(N^{-1})$$

and if $p\geqq 3$,

$$\begin{aligned}\Pr(\sqrt{N}[B_p(r_1)-B_p(\rho)-N^{-1}\tau(7-5p)/\{3kp(p-1)\}]/\tau<x)\\ =\Phi(x)+O(N^{-1})\end{aligned} \tag{2.3}$$

where $\tau=k[\{1+(p-1)\rho\}/(1-\rho)]^{(p-2)/(3p)}$, so that the remainder terms are of order N^{-1}. Thus, it can be shown that the transformations to achieve normality are $z_2(r_1)$ for $p=2$ and $B_p(r_1)$ for $p\geqq 3$.

A variance stabilizing transformation is obtained by solving the differential equation

$$\tau=k(1-\rho)\{1+(p-1)\rho\}\frac{df}{d\rho}=1$$

in (2.1) and the solution may be found to be the Fisher z transformation $z_p(r_1)$ given by (1.1). It is of interest to note that in the case $p=2$ this differential equation is exactly the same as that of (2.2).

Taking $f(r_1)=z_p(r_1)$ in (2.1) and choosing a correction factor suitably, we have

$$\begin{aligned}\Pr(\sqrt{N}[z_p(r_1)-z_p(\rho)-N^{-1}(7-5p)/\{3kp(p-1)\}]<x)\\ =\Phi(x)-N^{-1/2}[(2-p)/\{3kp(p-1)\}]x^2\phi(x)+O(N^{-1}),\end{aligned} \tag{2.4}$$

so that the term of $O(N^{-1/2})$ in the expansion cannot be made to vanish except for the case $p=2$ by $z_p(r_1)$.

It is worth pointing out that Fisher's z transformation satisfies, only in the case where $p=2$, the two requirements of the normalizing and variance stabilizing transformation simultaneously.

In the case where the mean vector μ is unrestricted, the maximum likelihood estimate of ρ has the form

$$r_2=\sum_{\alpha=1}^{N}\sum_{i\neq j}^{p}(X_{i\alpha}-\bar{X}_i)(X_{j\alpha}-\bar{X}_j)\Big/\Big\{(p-1)\sum_{\alpha=1}^{N}\sum_{i=1}^{p}(X_{i\alpha}-\bar{X}_i)^2\Big\}$$

where $\bar{X}_i=\sum_{\alpha} X_{i\alpha}/N$. An asymptotic expansion for the distribution of $f(r_2)$ is given by

$$\begin{aligned}&\Pr[\sqrt{n}\{f(r_2)-f(\rho)-c/n\}/\tau<x]\\&\quad=\Phi(x)-n^{-1/2}[(4-2p)/\{3kp(p-1)\}-c/\tau+a(\rho,f',f'')x^2]\phi(x)+O(n^{-1})\end{aligned}$$

where $n=N-1$ and τ, c, $\phi(x)$, $a(\rho, f', f'')$ are given in (2.1). By an argument similar to that used for r_1, we have that if $p=2$,

$$\Pr[\sqrt{n}\{z_2(r_2)-z_2(\rho)\}<x]=\Phi(x)+O(n^{-1})$$

and if $p\geqq 3$,

$$\Pr(\sqrt{n}[B_p(r_2)-B_p(\rho)-n^{-1}\tau(4-2p)/\{3kp(p-1)\}]/\tau<x)=\Phi(x)+O(n^{-1})$$

where $\tau=k[\{1+(p-1)\rho\}/(1-\rho)]^{(p-2)/(3p)}$ with k given in (2.1). Hence a similar result holds for the problem of finding transformations of the maximum likelihood estimate of ρ in the case where μ is unrestricted, except that the bias correction terms are slightly different.

The approach used here was also applied to derive normalizing and variance stabilizing transformations for r_1 in the case where the number of siblings varies among families. However such a transformation does not seem to be expressible as an elementary function.

3. Numerical comparisons

To examine whether the transformations given in the last section are efficient to achieve normality, we compare the accuracy of the approximations to the distribution of r_1 in terms of values of the probability integral. It is known (Olkin and Pratt [11]) that

$$N(N-1)^{-1}[\{1+(p-1)r_1\}/(1-r_1)][(1-\rho)/\{1+(p-1)\rho\}]$$

has the F distribution with $N-1$ and $(p-1)N$ degrees of freedom. We use as a measure of the approximation error: $\operatorname{Max}|\Pr(r_1<r_0;N,\rho)-A(r_0)|\times 10^4$ where $A(r_0)$ is obtained by using the approximate distri-

butions discussed in Section 2. Exact and approximate values are calculated at intervals of 0.001 between $-(p-1)^{-1}$ and 1.0 for various values of ρ, N and p. Table 1 gives an overall comparison of the four approximations concerning the maximum error. In the table, AD, Z I, Z II and BT are the notations standing for the following:

AD : the case that $\sqrt{N}(r_1-\rho)/[\sqrt{2}\{p(p-1)\}^{-1/2}(1-\rho)\{1+(p-1)\rho\}]$ is approximated by a normal variate with mean 0 and variance 1,

Z I : the case that $\sqrt{N}\{z_p(r_1)-z_p(\rho)\}$ is approximated by a normal variate with mean 0 and variance 1,

Z II : the case that the values of $\Pr(r_1<r_0)$ are approximated by using the leading term $\Phi(x)$ of (2.4),

BT : the case that the values of $\Pr(r_1<r_0)$ are approximated by (2.3).

Table 1. Comparison of maximum errors in approximating the values of $\Pr(r_1<r_0)$: Max |Exact value−Approximate value| × 10^4, where the maximum is over the values $r_0=-(p-1)^{-1}(0.001)1.0$.

ρ		$p=2$	$p=3$	$p=4$	$p=6$	$p=8$	$p=10$
				$N=25$			
0.1	AD	427	632	722	804	842	865
	Z I	422	637	729	812	852	875
	Z II	73	129	154	178	190	196
	BT		61	55	48	44	42
0.5	AD	546	671	745	820	857	878
	Z I	422	637	729	812	852	875
	Z II	73	129	154	178	190	196
	BT		61	55	48	44	42
0.9	AD	766	877	931	985	1012	1028
	Z I	422	637	729	812	852	875
	Z II	73	129	154	178	190	196
	BT		61	55	48	44	42

It may be seen from the table that the approximate distribution BT is the most accurate of these approximations for various values of ρ and p, and provides high accuracy over the whole domain of r_1. We observe that in comparison with Z I the approximation Z II can be markedly improved by adding the correction factor.

Fig. 1 shows the error between approximate and exact values of $\Pr(r_1<r_0)$, for which the approximations (2.3) and (2.4) neglected the term of $O(N^{-1/2})$ are used.

From the results presented in Table 1 and Fig. 1, it is interesting to note that the maximum error of BT decreases steadily as p increases, while other approximate distributions depart from normality. The curve in the case $p=2$ of Fig. 1. (b) based on $z_2(r_1)$ has the same pattern as ones in Fig. 1. (a) based on $B_p(r_1)$. It must be remembered that the transformation of the form of $z_p(r_1)$ yields a normalization only in the case where $p=2$.

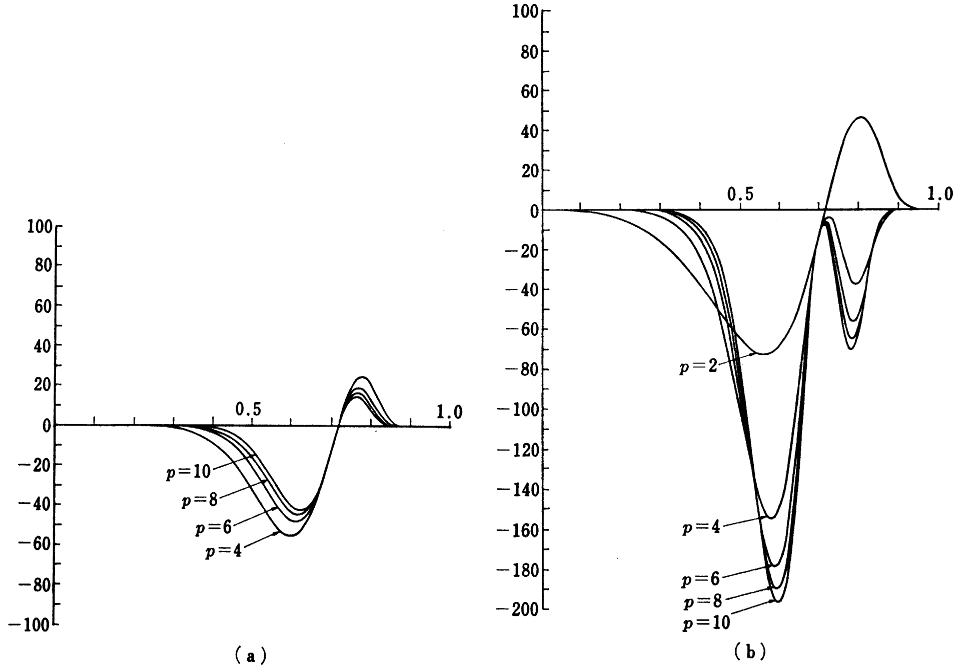

Fig. 1. Errors in approximating the values of $\Pr(r_1<r_0)$ by using (2.3); (a), and the leading term of (2.4); (b), for $N=25$ and $\rho=0.7$: Error=(Approximate value−Exact value)

It may be concluded that the transformation $B_p(r_1)$ derived here is successful in normalizing the distribution of r_1, although the effect of the variance stabilizing transformation is lost. Similar comparisons were made for the approximate distributions of r_2. In consequence we found the results to be unchanged essentially.

Remark. It follows from the result of Olkin and Pratt [11] that the distributions of $B_p(r_1)/B_p(\rho)$ and $z_p(r_1)-z_p(\rho)$ are independent of the population intraclass correlation ρ. This shows immediately that the approximations ZI, ZII and BT are also independent of ρ and that $z_p(r_1)$ yields the variance stabilizing transformation. Hence the maximum errors for the approximations ZI, ZII and BT are independent of the values of ρ unlike the case of the sample correlation coefficient in a bivariate normal sample.

Acknowledgements

The author would like to thank the referee for helpful comments and suggestions, particularly with regard to Remark in Section 3, and for pointing out some references. He also thanks Mrs. F. Sato for

carrying out the computer analyses.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.*, **6**, 434-451.

[2] Borges, R. (1971). Derivation of normalizing transformations with an error of order $1/n$, *Sankhyā* Ser. A, **33**, 441-460.

[3] Efron, B. (1982). Transformation theory: How normal is a family of distributions? *Ann. Statist.*, **10**, 323-339.

[4] Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample, *Metron*, **1**, 1-32.

[5] Fisher, R. A. (1958). *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh.

[6] Fujikoshi, Y. (1980). Asymptotic expansions for the distributions of the sample roots under nonnormality, *Biometrika*, **67**, 45-51.

[7] Hougaard, P. (1982). Parametrizations of non-linear models, *J. R. Statist. Soc.*, B, **44**, 244-252.

[8] Konishi, S. (1978). An approximation to the distribution of the sample correlation coefficient, *Biometrika*, **65**, 654-656.

[9] Konishi, S. (1981). Normalizing transformations of some statistics in multivariate analysis, *Biometrika*, **68**, 647-651.

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

[11] Olkin, I. and Pratt, J. W. (1958). Unbiased estimation of certain correlation coefficients, *Ann. Math. Statist.*, **29**, 201-211.

[12] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., John Wiley and Sons, New York.