Testing whether a population correlation coefficient corrected for attenuation is perfect

GORDON RAE

New University of Ulster
Coleraine, Londonderry BT52 1SA, Northern Ireland

The statistical problem of deciding whether the disattenuated correlation (ρ_{tt}) of two measures may be assumed to be perfect is of obvious practical significance. It arises, for example, when the validity of a uniprocess theory of retroactive and proactive inhibition is tested or when individual differences in auditory and visual memory are investigated (see, for example, Jensen, 1971).

The present program enables the investigator to carry out a number of procedures developed for this purpose (see below). With the exception of Kristof's (1973) technique, it is assumed that each measure can be divided into two parallel halves or that two parallel forms for each measure are available.

Lord's (1957) Procedure. This procedure involves computing a likelihood ratio statistic, χ^2_1 , which is distributed as chi square with 1 degree of freedom. If χ^2_1 is significant then the null hypothesis $H_0: \rho_{tt} = 1.0$ is rejected.

With the exception of $\tilde{\sigma}_{12}$, $\tilde{\sigma}_{34}$, and $\tilde{\sigma}_{13}$ (the maximum likelihood estimates of the unknown parameters σ_{12} , σ_{34} , and σ_{13} under the null hypothesis), all values in the ratio statistic can be readily obtained from the sample variance-covariance matrix (s_{ij}) ($i, j = 1, 2, 3, 4$) using Lord's (1957) Equations 10-14 and Equations 27 and 48. Lord suggests that $\tilde{\sigma}_{12}$ and $\tilde{\sigma}_{34}$ can be found by plotting Equations 40 and 41 and finding their intersection. A larger and larger scale can then be used in successive plots in order to locate the intersection accurately. Unfortunately, such a procedure does not readily lend itself to computer programming, and the following approach was adopted instead.

If we denote $\tilde{\sigma}_{12}$ and $\tilde{\sigma}_{34}$ by x and y , respectively, then Lord's (1957) Equations 40 and 41 may be written in the form $f(x, y) = 0$ and $g(x, y) = 0$. In order to solve these nonlinear simultaneous equations, we can use an extension of the Newton-Raphson method. This states that, subject to certain conditions, if x_i and y_i are approximations to the roots of these equations, better approximations are given by $x_{i+1} = x_i + \epsilon_i$ and $y_{i+1} = y_i + \eta_i$, where $\epsilon_i [\partial f(x_i y_i) / \partial x] + \eta_i [\partial f(x_i y_i) / \partial y] = -f(x_i y_i)$ and $\epsilon_i [\partial g(x_i y_i) / \partial x] + \eta_i [\partial g(x_i y_i) / \partial y] = -g(x_i y_i)$. As first approximations to the roots, we may take $x_0 = s_{12}$ and $y_0 = s_{34}$ (the sample covariances). Iterations can

then be continued until $|x_{i+1} - x_i|$ and $|y_{i+1} - y_i|$ are both less than some specified value, say 10^{-4} . When the values of $\tilde{\sigma}_{12}$ and $\tilde{\sigma}_{34}$ have been determined, the value of $\tilde{\sigma}_{13}$ may then be obtained from Lord's Equation 26.

In order to check the Newton-Raphson procedure, Lord's (1957) sample covariance matrix (p. 210) was used as input data. After only four iterations, the values of $\tilde{\sigma}_{12}$ and $\tilde{\sigma}_{34}$ obtained (51.6404 and 71.2683, respectively) were virtually identical to those found by Lord using the more cumbersome graphical method.

Note that under certain circumstances an iterative solution may break down. However, Monte Carlo studies carried out by the present author suggest that, when s_{12} and s_{34} are used as initial approximations, this very rarely occurs in practice.

McNemar's (1958) Procedure. Using an analysis of variance approach, McNemar (1958) has developed a number of statistics for testing the hypothesis that the disattenuated correlation coefficient in the population is perfect. According to McNemar, the most useful statistic for this purpose is the mean square ratio (MSR), given by his Equation 7. If the MSR obtained is greater than F_p , where F_p denotes the $1 - p$ th percentile of the F distribution with $(N - 1)/2(N - 1)$ degrees of freedom, then the null hypothesis $H_0: \rho_{tt} = 1.0$ is rejected at significance level $< p$.

Forsyth and Feldt's (1970) Modification. As Forsyth and Feldt (1970) point out, McNemar's (1958) MSR statistic is not that of a central F with $N - 1$ and $2(N - 1)$ degrees of freedom unless the two tests are equally reliable. Forsyth and Feldt suggest that if the equal reliability assumption is not met, the MSR obtained should be divided by a constant, K_H , which is a function of the reliability coefficients of the tests and the correlation between them. This modified test provides better control of Type I errors when the reliabilities are not equal, although it is not entirely satisfactory when the reliabilities are markedly different.

Forsyth and Feldt's (1969) Procedure. This procedure is based on the fact that, for suitably large samples, the sampling distribution of the disattenuated correlation coefficient is approximately normal in form and that a satisfactory estimate of the standard error can be obtained by evaluating Kelley's (1947) formula with the hypothesized value of ρ_{tt} replacing the actual sample value. Thus, in order to reject the hypothesis that $\rho_{tt} = 1.0$, the sample disattenuated correlation coefficient must fall below $1.0 - Z_\alpha \text{estSE}$, where Z_α denotes the value of the normal deviate, with α of the area below it, and estSE denotes the estimate of the standard error. Note that the critical region is placed entirely at the lower end of the hypothesized sampling distribution, since $H_a: \rho_{tt} < 1$.

Lord's (1973) Procedure. This technique is based on the construction of a confidence interval for the coefficients of a linear relation between true scores. It involves determining the matrix M , defined by $M = A - F_p W$, where A denotes the 2 by 2 matrix of among-persons sums of squares and cross products, W denotes the corresponding 2 by 2 within-persons matrix, and F_p denotes the 1 - pth percentile of the F distribution with N/N degrees of freedom. If the matrix M is positive definite, then the null hypothesis is rejected at significance level $< p$. Otherwise, the hypothesis is not rejected.

Since Matrix W cannot normally be constructed from the sample variance-covariance matrix, this technique is omitted unless the input data consist of raw scores.

Kristof's (1973) Procedure. This technique is designed to test the null hypothesis (H_0^*) that there is an unspecified linear relation between the true scores on the two measures. Kristof (1973) considers three different cases, but since any set of data satisfying the assumptions of Cases 1 and 2 may be subjected to the test developed for Case 3, only the last is included in the present program. Case 3 makes fewer assumptions than any other technique used in the program, since corresponding forms of each measure need not be parallel and the errors of measurement need not all covary zero.

The procedure involves partitioning the sample variance-covariance matrix into Submatrices V_{11} , V_{22} , V_{12} , and V'_{12} . It is assumed that Matrix V_{12} is positive definite; otherwise the null hypothesis could not be rejected at any significance level. Since V_{12} will not normally be symmetric, it is replaced by the symmetric matrix $(V_{12} + V'_{12})/2$, and the canonical decomposition of the latter matrix is then obtained. Equation 12 is then used to generate the elements of two other matrices, A and B , which provide the coefficients for Kristof's (1973) quartic equation in $\text{ctg } \phi$ (Kristof's Equation 26). In the computer program, the roots of this equation (both real and imaginary) are found using the Lin-Bairstow iterative method (see McCalla, 1967, pp. 115-129). For each of the real roots, the quantity r^* is found using Kristof's Equation 23. The smallest (positive) value of r^* is used to test H_0^* . This minimum can be tested using the transformation $t^* = r^*[(N-2)/(1-r^{*2})]^{1/2}$. If t^* is greater than the critical value of t using a one-tailed test (level α , $df = N-2$), then the null hypothesis is rejected.

Although this and Lord's (1973) procedure both tend to yield conservative results, they are appropriate for any sample size.

Input. Raw scores or the sample variance-covariance may be used as input data. In addition, the user must supply the sample size (N), the maximum number of iterations to be performed in obtaining Lord's (1957)

estimates $\tilde{\sigma}_{12}$ and $\tilde{\sigma}_{34}$, the absolute difference between successive approximations that will be used to indicate that convergence has been achieved, and the 95th and 99th percentiles of the F distribution with N/N degrees of freedom. (The last are necessary for Lord's, 1973, procedure.)

Output. The following are given as printed output: Lord's (1957) disattenuated correlation coefficient, his likelihood ratio statistic (χ^2_1) and χ_1 ; McNemar's (1958) MSR statistic, the corresponding degrees of freedom, and Forsyth and Feldt's (1970) modified statistic MSR/ K_H ; Forsyth and Feldt's (1969) estimate of the standard error of the disattenuated correlation coefficient and the critical regions for the .05 and .01 significance levels; and the roots of Kristof's (1973) quartic equation (real and imaginary) and the values of r^* and t^* corresponding to the real roots. (The smallest value of r^* should be used to test the null hypothesis.) If raw scores are used as input data, the sample variance-covariance matrix, the sample correlation matrix, and a statement about whether one should or should not accept the null hypothesis using Lord's (1973) procedure are also included as printed output.

Sample output, using Lord's (1957, p. 210) sample variance-covariance matrix as input data, is shown in the appendix.

Computer and Language. The program is written in FORTRAN IV and was prepared and tested on the New University of Ulster VAX-11 computer.

Availability. A source listing may be obtained without charge from Gordon Rae, Education Centre, New University of Ulster, Coleraine, Northern Ireland.

REFERENCES

- FORSYTH, R. A., & FELDT, L. S. An investigation of empirical sampling distributions of correlation coefficients corrected for attenuation. *Educational and Psychological Measurement*, 1969, 29, 61-71.
- FORSYTH, R. A., & FELDT, L. S. Some theoretical and empirical results related to McNemar's test that the population correlation coefficient corrected for attenuation equals 1.0. *American Educational Research Journal*, 1970, 7, 197-207.
- JENSEN, A. R. Individual differences in visual and auditory memory. *Journal of Educational Psychology*, 1971, 62, 123-131.
- KELLEY, T. L. *Fundamentals of statistics*. Cambridge: Harvard University Press, 1947.
- KRISTOF, W. Testing a linear relation between true scores of two measures. *Psychometrika*, 1973, 38, 101-111.
- LORD, F. M. A significance test for the hypothesis that two variables measure the same trait except for errors of measurement. *Psychometrika*, 1957, 22, 207-220.
- LORD, F. M. Testing if two measuring procedures measure the same dimension. *Psychological Bulletin*, 1973, 79, 71-72.
- MCCALLA, T. R. *Introduction to numerical methods and FORTRAN programming*. New York: Wiley, 1967.
- MCMENAR, Q. Attenuation and interaction. *Psychometrika*, 1958, 23, 259-266.

Appendix

LORD'S SIGMA12 TILDE= 51.640472
 LORD'S SIGMA34 TILDE= 71.268288
 NUMBER OF ITERATIONS= 4
 DISATTENUATED CORR OF LURD= 0.898643
 CHI SQUARE= 35.400421
 CHI= 5.949825
 MCNEMAR'S F RATIO= 1.512000
 DEGREES OF FREEDOM FOR NUMERATOR= 648.
 DEGREES OF FREEDOM FOR DENOMINATOR=1296.
 FORSYTH AND FELDT'S MODIFIED F= 1.497977
 FORSYTH AND FELDT'S DISATTENUATED CORR= 0.898637
 FORSYTH AND FELDT'S STANDARD ERROR= 0.01401163
 CRITICAL REGION AT .05 LEVEL INCLUDES VALUES LESS THAN 0.976951
 CRITICAL REGION AT .01 LEVEL INCLUDES VALUES LESS THAN 0.967353
 FOR KRISTOF'S CASE 3 ROOTS OF QUARTIC EQUATION ARE:
 (REAL PART) (IMAGINARY PART)
 -0.032237422 2.125279427
 -0.032237422 -2.125279427
 0.019372940 0.000000000
 -51.867149353 0.000000000
 VALUES OF R* AND T* CORRESPONDING TO THE REAL
 ROOTS ARE:

R*	T*	DEGREES OF FREEDOM
0.183873	4.758153	647.
0.835061	38.608585	647.

(Accepted for publication June 1, 1982.)